

A CLT for an improved subspace estimator with observations of increasing dimensions

P. Vallet⁽¹⁾, X. Mestre⁽²⁾, P. Loubaton⁽³⁾

Abstract

This paper deals with subspace estimation in the small sample size regime, where the number of samples is comparable in magnitude with the observation dimension. The traditional estimators, mostly based on the sample correlation matrix, are known to perform well as long as the number of available samples is much larger than the observation dimension. However, in the small sample size regime, the performance degrades. Recently, based on random matrix theory results, a new subspace estimator was introduced, which was shown to be consistent in the asymptotic regime where the number of samples and the observation dimension converge to infinity at the same rate. In practice, this estimator outperforms the traditional ones even for certain scenarios where the observation dimension is small and of the same order of magnitude as the number of samples. In this paper, we address a performance analysis of this recent estimator, by proving a central limit theorem in the above asymptotic regime. We propose an accurate approximation of the mean square error, which can be evaluated numerically.

Contents

1	Introduction	2
1.1	Motivation	2
1.2	Notations	3
2	Asymptotic behaviour of the sample eigenvalues	4
2.1	The asymptotic spectral distribution	4
2.2	Useful quantities and bounds	6
2.3	Separation of the sample eigenvalues	7
2.4	The spiked model case: fixed rank	9
2.5	Contour integrals	12
3	Noise subspace estimation	12
3.1	Consistent estimation	12
3.2	CLT	14
3.2.1	The main result	14
3.2.2	Discussions and numerical examples	15
3.2.3	CLT for the traditional noise subspace estimate	16
4	Proof of theorem 1	17
4.1	Regularization and confinement of the eigenvalues	17
4.2	The differential equation	18
4.3	Asymptotics of $\vartheta_N(k, \ell)$	21

⁽¹⁾ Laboratoire de l'Intégration du Matériau au Système (CNRS, Univ. Bordeaux, Bordeaux INP), 351, Cours de la Libération 33405 Talence (France), pascal.vallet@bordeaux-inp.fr

⁽²⁾ Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Av. Carl Friedrich Gauss 08860 Castelldefels, Barcelona (Spain), xavier.mestre@cttc.cat

⁽³⁾ Université Paris-Est / Laboratoire d'Informatique Gaspard Monge (CNRS UMR 8049), 5 Bd. Descartes 77454 Marne-la-Vallée (France), loubaton@univ-mlv.fr

This work was partially supported by the French programs GDR ISIS/GRETSI "Jeunes Chercheurs" and Project ANR-12-MONU-0003 DIONISOS Project.

4.4	Asymptotics of $\vartheta_N(k, l)$	23
4.5	Solution to the differential equation	24
5	Appendix	25
5.1	Proof of formula (66) and proposition 1	25
5.1.1	Proof of formula (66)	27
5.1.2	Proof of proposition 1	29
5.2	Proof of lemma 1	31
5.3	Proof of lemma 3	33

1 Introduction

1.1 Motivation

The problem of subspace estimation, i.e. estimating the eigenspaces of the correlation matrix of a certain multivariate time series of dimension M , available from a set of N noisy observations, is an important problem in statistical signal processing, and covers several topics such as Direction of Arrival (DoA) estimation [14], multiuser detection in Code Division Multiple Access (CDMA) [8], chirp parameter estimation [17] or beamforming [3]. Let us consider an complex M -variate time series $(\mathbf{y}_n)_{n \geq 1}$, following a "signal plus noise" model

$$\mathbf{y}_n = \mathbf{s}_n + \mathbf{v}_n,$$

where \mathbf{s}_n corresponds to a signal part and \mathbf{v}_n to a noise part, and assume that N observations $\mathbf{y}_1, \dots, \mathbf{y}_N$ are collected and stacked in the $M \times N$ matrix

$$\mathbf{Y}_N = [\mathbf{y}_1, \dots, \mathbf{y}_N] = \mathbf{S}_N + \mathbf{V}_N,$$

with $\mathbf{S}_N = [\mathbf{s}_1, \dots, \mathbf{s}_N]$ and $\mathbf{V}_N = [\mathbf{v}_1, \dots, \mathbf{v}_N]$. In many applications, the signals $(\mathbf{s}_n)_{n \geq 0}$ are moreover constrained to a subspace of dimension K less than M and the matrix \mathbf{S}_N is full rank K . The subspace estimation problem consists in estimating the column space of \mathbf{S}_N called the "signal subspace", of dimension K (or equivalently its orthogonal complement called the "noise subspace" of dimension $M - K$) from the observation matrix \mathbf{Y}_N .

The usual way of estimating the signal or noise subspaces consists in estimating their orthogonal projection matrices. The estimation is performed most of the time by using the so-called sample correlation matrix (SCM) of the observations

$$\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N} = \frac{1}{N} \sum_{n=1}^N \mathbf{y}_n \mathbf{y}_n^*,$$

and these projections are directly estimated by considering their sample estimates, i.e. by considering the corresponding orthogonal projection matrix of the SCM. For example, the noise subspace projection matrix $\mathbf{\Pi}_N$, i.e. the orthogonal projection matrix onto the kernel of $\mathbf{S}_N \mathbf{S}_N^*$, is traditionally estimated by $\hat{\mathbf{\Pi}}_N$, the orthogonal projection matrix onto the eigenspace associated with the $M - K$ smallest eigenvalues of $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}$.

These sample estimators are known to perform well when the number of available samples N is much larger than the observation dimension M , in particular because the SCM is a good estimator of the true correlation matrix of the observations. Indeed, in the asymptotic regime where M is constant and N converges to infinity, under some technical conditions, the law of large numbers ensures that

$$\left\| \hat{\mathbf{\Pi}}_N - \mathbf{\Pi}_N \right\| \rightarrow 0$$

almost surely as $N \rightarrow \infty$, i.e. the sample projection matrices are consistent estimators of the true ones. These sample estimators have been also characterized in terms of Central Limit Theorems (CLT) in the previous asymptotic regime, and several accurate approximations of the Mean Square Error (MSE) have been obtained, see e.g. Anderson [1], Stoica [15], and the references therein. However, it may exist some situations where obtaining such an amount of samples is not conceivable, for example in situations where

the signals are stationary only for a short period of time, or simply if the observation dimension is large. As a consequence, in the low sample size regime where M and N are of the same order of magnitude, the performance of the sample subspace estimators severely degrades, essentially because the SCM does not estimate properly the true correlation matrix.

In this context, based on recent results in random matrix theory, a new subspace estimator was proposed by Mestre [11], in the case where $(\mathbf{s}_n)_{n \geq 0}$ and $(\mathbf{v}_n)_{n \geq 0}$ are modeled as two independent zero-mean Gaussian stationary processes, temporally uncorrelated, with the signal correlation matrix $\mathbf{R}_s = \mathbb{E}[\mathbf{s}_n \mathbf{s}_n^*]$ being rank K and the noise covariance being equal to $\sigma^2 \mathbf{I}$, where $\sigma > 0$ and \mathbf{I} is the $M \times M$ identity matrix, i.e. $(\mathbf{y}_n)_{n \geq 0}$ can be modeled equivalently as

$$\mathbf{y}_n = (\mathbf{R}_s + \sigma^2 \mathbf{I})^{1/2} \mathbf{x}_n, \quad (1)$$

with $(\mathbf{x}_n)_{n \geq 0}$ a standard spatially and temporally white Gaussian process. Later Vallet et al. [16] obtained, using the same approach, a different estimator in the more general situation where the signals $(\mathbf{s}_n)_{n \geq 0}$ are considered as unknown deterministic. The estimators of [11] and [16] were shown to be consistent in the asymptotic regime where both the observation dimension M and the number of samples N converge to infinity in such a way that the ratio $\frac{M}{N}$ converge to a positive constant. Moreover, these estimators do not assume any particular assumption on the behaviour of the rank K , which may also converge to infinity with M, N . In practice, these estimators outperform the traditional ones, when M, N are of the same order of magnitude. Based on these results, an application to DoA estimation of K source signals impinging on an array of M sensors was proposed, and an improved subspace DoA estimator called G-MUSIC (Generalized Multiple Signal Classification) was built, which was shown to numerically outperform the traditional MUSIC estimator, for realistic values of M, N . This DoA estimator was also shown to be consistent in Hachem et al. [7].

Recently, Hachem et al. [6] proposed an analysis of the subspace estimator [16], in terms of a Central Limit Theorem (CLT), in the previous asymptotic regime where M, N converge to infinity at the same rate, and by assuming that the rank K is fixed. In practice, these results are accurate as long as the rank K remains small compared to M, N . However, when the rank K is of the same order of magnitude than the dimension M and N , the corresponding results do not predict anymore the behaviour of the subspace estimator [16] and the results of [6] are not very accurate.

In this paper,¹ we propose to extend the analysis of [6] regardless the behaviour of the rank K , which may increase with M, N . For that purpose, we use a different approach and prove a Central Limit Theorem (CLT) in the previous asymptotic regime. We also provide an explicit expression for the Mean Square Error (MSE) which can be easily evaluated numerically. Numerical examples confirm the validity of the results.

The paper is organized as follows. In the remainder of section 1, we introduce formally the model of signals used in the paper, and recall some basic results from random matrix theory, necessary for the next sections. In section 3, we introduce the subspace estimator of [16] and provide the main result of the paper, namely a CLT for this estimator, as well as numerical illustrations. Sections 4 and 5 contain the proofs of the results.

1.2 Notations

We introduce here the main notations used throughout the paper.

The sets \mathbb{R} , \mathbb{R}^+ and \mathbb{N} (resp. \mathbb{N}^*) will respectively represent the real numbers, the non-negative numbers and the non-negative integers (resp. the positive integers). \mathbb{C} will be the set of complex numbers, and for $z \in \mathbb{C}$, $\text{Re}(z)$, $\text{Im}(z)$ and z^* will stand for the real part, the imaginary part and the complex conjugate. i will be the imaginary unit and we will also use the set $\mathbb{C}^+ = \{z \in \mathbb{C} : \text{Im}(z) > 0\}$. The indicator of a set $\mathcal{E} \subset \mathbb{R}$ is denoted $\mathbb{1}_{\mathcal{E}}$, $\partial \mathcal{E}$ and $\text{Int}(\mathcal{E})$ will denote the boundary and interior of \mathcal{E} .

For a real-valued function φ defined on \mathbb{R} , $\text{supp}(\varphi)$ will represent the support of φ , and $\mathcal{C}_c^\infty(\mathbb{R}, \mathcal{E})$ will the set of smooth compactly supported functions defined on \mathbb{R} , taking values in some set $\mathcal{E} \subset \mathbb{R}$.

Matrices (respectively vectors) are denoted by boldfaced capital (respectively boldfaced lower case) letters. For a complex matrix \mathbf{A} , we denote by \mathbf{A}^T , \mathbf{A}^* its transpose and its conjugate transpose, and by $\text{tr}(\mathbf{A})$ and $\|\mathbf{A}\|$ its trace and spectral norm. The identity matrix will be \mathbf{I} . \mathbf{e}_n will refer to a vector having all its components equal to 0 except the n -th equals to 1.

¹The material of this paper was partly presented in the conference paper [12].

The real normal distribution with mean m and variance σ^2 is denoted $\mathcal{N}_{\mathbb{R}}(\alpha, \sigma^2)$ and the multivariate normal distribution in \mathbb{R}^k , with mean \mathbf{m} and covariance $\mathbf{\Gamma}$ is denoted in the same way $\mathcal{N}_{\mathbb{R}^k}(\mathbf{m}, \mathbf{\Gamma})$. We will say that a complex random variable $Z = X + iY$ follow the distribution $\mathcal{N}_{\mathbb{C}}(\alpha + i\beta, \sigma^2)$ if X and Y are independent with respective distributions $\mathcal{N}_{\mathbb{R}}(\alpha, \frac{\sigma^2}{2})$ and $\mathcal{N}_{\mathbb{R}}(\beta, \frac{\sigma^2}{2})$. The expectation and variance of a complex random variable Z will be denoted $\mathbb{E}[Z]$ and $\mathbb{V}[Z]$. The support of a probability measure μ will be denoted $\text{supp}(\mu)$. For a sequence of random variables $(X_n)_{n \in \mathbb{N}}$ and a random variable X , we write

$$X_n \xrightarrow[n \rightarrow \infty]{a.s.} X \text{ and } X_n \xrightarrow[n \rightarrow \infty]{\mathcal{D}} X$$

when X_n converges respectively with probability one and in distribution to X . Finally, $X_n = o_{\mathbb{P}}(1)$ will stand for the convergence of X_n to 0 in probability, and $X_n = \mathcal{O}_{\mathbb{P}}(1)$ will stand for boundedness in probability (tightness).

Some other special notations may be used at some very localized parts in the paper, and will be introduced in the text.

2 Asymptotic behaviour of the sample eigenvalues

In this section, we present some basic results from random matrix theory, describing the behaviour of the eigenvalues of the SCM $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}$, in the asymptotic regime where M, N converge to infinity such that $\frac{M}{N} \rightarrow c > 0$. These results will be required to introduce the improved subspace estimator of [16]. To that end, we will work with the following random matrix model, referred to as "Information plus Noise" in the literature. We consider $M, N, K \in \mathbb{N}^*$ such that $K < M < N$ and $M = M(N)$, $K = K(N)$ are functions of N satisfying $c_N = \frac{M}{N} \rightarrow c \in (0, 1)$ as $N \rightarrow \infty$. For each $N \in \mathbb{N}^*$, we consider the $M \times N$ random matrix $\mathbf{\Sigma}_N$, defined by

$$\mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N, \tag{2}$$

with

- \mathbf{B}_N a rank K deterministic matrix satisfying $\sup_N \|\mathbf{B}_N\| < \infty$,
- \mathbf{W}_N having i.i.d. entries $W_{i,j} \sim \mathcal{N}_{\mathbb{C}}\left(0, \frac{\sigma^2}{N}\right)$.

We denote by $\lambda_{1,N} > \dots > \lambda_{K,N} > \lambda_{K+1,N} = \dots = \lambda_{M,N} = 0$ the eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ (the non-zero eigenvalues are assumed to have multiplicity one for simplicity), and by $\mathbf{u}_{1,N}, \dots, \mathbf{u}_{M,N}$ the respective unit norm eigenvectors. Equivalently, $\hat{\lambda}_{1,N} \geq \dots \geq \hat{\lambda}_{M,N}$ are the eigenvalues of the matrix $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$ and $\hat{\mathbf{u}}_{1,N}, \dots, \hat{\mathbf{u}}_{M,N}$ the respective unit norm eigenvectors.

2.1 The asymptotic spectral distribution

Let $\hat{\mu}_N = \frac{1}{M} \sum_{k=1}^M \delta_{\hat{\lambda}_{k,N}}$ the empirical spectral measure of the matrix $\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*$, with δ_x the Dirac measure at point x . From Dozier & Silverstein [5] [4], there exists a deterministic probability measure μ_N , with support $\text{supp}(\mu_N) \subset \mathbb{R}^+$, such that w.p.1.,

$$\hat{\mu}_N - \mu_N \xrightarrow[N \rightarrow \infty]{w} 0,$$

where "w" stands for the weak convergence. Equivalently the Stieltjes transform $\hat{m}_N(z)$ of $\hat{\mu}_N$, defined by

$$\hat{m}_N(z) = \int_{\mathbb{R}} \frac{d\hat{\mu}_N(\lambda)}{\lambda - z} = \frac{1}{M} \text{tr } \mathbf{Q}_N(z)$$

where $\mathbf{Q}_N(z) = (\mathbf{\Sigma}_N \mathbf{\Sigma}_N^* - z\mathbf{I})^{-1}$ satisfies for all $z \in \mathbb{C} \setminus \mathbb{R}$

$$\hat{m}_N(z) - m_N(z) \xrightarrow[N \rightarrow \infty]{a.s.} 0,$$

where $m_N(z) = \int_{\mathbb{R}} \frac{d\mu_N(\lambda)}{\lambda - z}$ is the Stieltjes transform of μ_N , which satisfies the equation

$$m_N(z) = \frac{1}{M} \text{tr } \mathbf{T}_N(z), \quad (3)$$

for all $z \in \mathbb{C} \setminus \mathbb{R}$, where the matrix $\mathbf{T}_N(z)$ is defined by

$$\mathbf{T}_N(z) = \left(\frac{\mathbf{B}_N \mathbf{B}_N^*}{1 + \sigma^2 c_N m_N(z)} - z(1 + \sigma^2 c_N m_N(z)) \mathbf{I} + \sigma^2(1 - c_N) \mathbf{I} \right)^{-1}.$$

Moreover, $m_N(z)$ can be further continuously extended to the real axis when $z \in \mathbb{C}^+ \rightarrow x \in \mathbb{R}$, and we denote the limit $m_N(x)$. Defined in this way, $x \mapsto m_N(x)$ is continuous on \mathbb{R} , continuously differentiable on $\mathbb{R} \setminus \partial \text{supp}(\mu_N)$ and still satisfies the equation (3) for $x \in \mathbb{R} \setminus \partial \text{supp}(\mu_N)$.

We now recall the characterization of the support of μ_N provided in [16]. Define the function $w_N(z)$ by

$$w_N(z) = z(1 + \sigma^2 c_N m_N(z))^2 - \sigma^2(1 - c_N)(1 + \sigma^2 c_N m_N(z)), \quad (4)$$

The main equation (3) can be expressed in terms of an equation in $w_N(z)$, i.e.

$$z = \phi_N(w_N(z)), \quad (5)$$

where

$$\phi_N(w) = w(1 - \sigma^2 c_N f_N(w))^2 + \sigma^2(1 - c_N)(1 - \sigma^2 c_N f_N(w)) \quad (6)$$

and

$$f_N(w) = \frac{1}{M} \text{tr } (\mathbf{B}_N \mathbf{B}_N^* - w \mathbf{I})^{-1}. \quad (7)$$

Starting from the properties that w_N is real and increasing on $\mathbb{R} \setminus \text{supp}(\mu_N)$ and $w_N(x) \in \mathbb{C}^+$ for $x \in \text{supp}(\mu_N)$, [16] characterized $w_N(x)$ among the set of all solutions of the polynomial equation $\phi_N(w) = x$ (which has degree $2K+2$), for $x \in \mathbb{R}$, and showed that ϕ_N admits $2Q$ ($1 \leq Q \leq K+1$) positive local extrema ² $0 < x_{1,N}^- < x_{1,N}^+ < \dots < x_{Q,N}^- < x_{Q,N}^+$ whose preimages are

$$w_N(x_{1,N}^-) < 0 < w_N(x_{1,N}^+) < \dots < w_N(x_{Q,N}^-) < w_N(x_{Q,N}^+). \quad (8)$$

Moreover, we always have $w_N(x_{Q,N}^+) > \lambda_{1,N}$, and if $Q > 1$, it turns out that for each $q = 1, \dots, Q-1$, there exists $k \in \{0, \dots, K\}$ such that

$$w_N(x_{q,N}^+), w_N(x_{q+1,N}^-) \in (\lambda_{k,N}, \lambda_{k+1,N}).$$

By differentiating (5) on both sides, we find $\phi'_N(w_N(x)) > 0$ for all $x \in \mathbb{R} \setminus \text{supp}(\mu_N)$. Finally, by showing that $\text{Im}(w_N(x)) = 0$ for $x \in \mathbb{R} \setminus \bigcup_{q=1}^Q [x_{q,N}^-, x_{q,N}^+]$ and $\text{Im}(w_N(x)) > 0$ for $x \in \bigcup_{q=1}^Q [x_{q,N}^-, x_{q,N}^+]$, [16] concludes that the support of μ_N is given by the union

$$\text{supp}(\mu_N) = \bigcup_{q=1}^Q [x_{q,N}^-, x_{q,N}^+], \quad (9)$$

where the intervals $[x_{q,N}^-, x_{q,N}^+]$ are called "clusters".

A typical illustration of function $\phi_N(w)$ for $w \in \mathbb{R}$ is given in figure 1.

²Note that $Q = Q(N)$ is a function of N

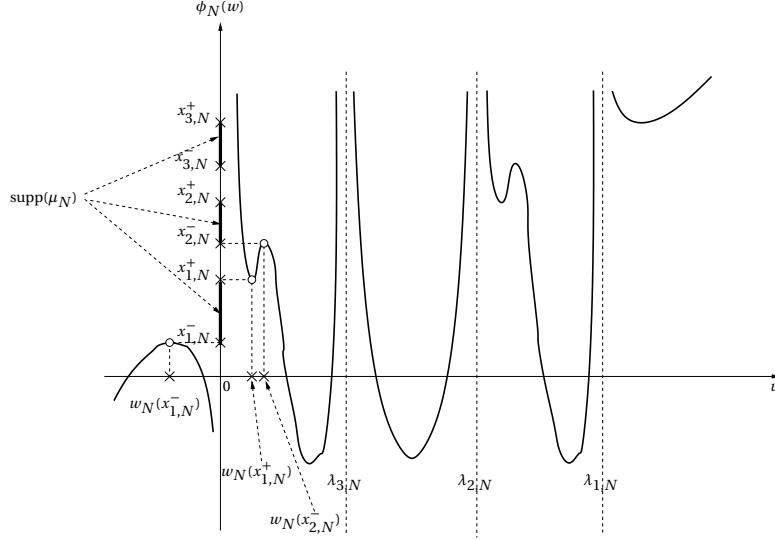


Figure 1: Typical example of the function ϕ_N , for $K = 3$. For $x \in (x_{q,N}^+, x_{q+1,N}^-)$ with $q = 1, \dots, Q-1$, the equation $\phi_N(w) = x$ admits $2K + 2$ real solutions and $w_N(x)$ is the unique solution in the interval $(w_N(x_{q,N}^+), w_N(x_{q+1,N}^-))$, and for $x < x_{1,N}^-$ (resp. $x > x_{Q,N}^+$), $w_N(x)$ is the unique solution in $(-\infty, w_N(x_{1,N}^-))$ (resp. $(w_N(x_{Q,N}^+), \infty)$). Conversely, for $x \in \text{supp}(\mu_N)$, the equation $\phi_N(w) = x$ admits $2K$ real solutions plus two complex conjugate solutions, and $w_N(x)$ coincides with the solution having positive imaginary part.

2.2 Useful quantities and bounds

We now introduce a few bounds which will be of constant use for the derivation of the main results of the paper. Let us define

$$\tilde{m}_N(z) = c_N m_N(z) - \frac{1 - c_N}{z}, \quad (10)$$

which corresponds to the Stieltjes transform of the probability measure $c_N \mu_N + (1 - c_N) \delta_0$. It can be shown that $\tilde{m}_N(z) = \frac{1}{N} \text{tr } \tilde{\mathbf{T}}_N(z)$, with

$$\tilde{\mathbf{T}}_N(z) = \left(\frac{\mathbf{B}_N^* \mathbf{B}_N}{1 + \sigma^2 \tilde{m}_N(z)} - z(1 + \sigma^2 c_N m_N(z)) \mathbf{I} \right)^{-1}, \quad (11)$$

and note that $w_N(z)$ defined in (4) can be written as

$$w_N(z) = z (1 + \sigma^2 c_N m_N(z)) (1 + \sigma^2 \tilde{m}_N(z)).$$

The proof of the following bounds can be found in [9], [16] and [7]: matrices $\mathbf{T}_N(z)$ and $\tilde{\mathbf{T}}_N(z)$ satisfy

$$\|\mathbf{T}_N(z)\| \leq \frac{C}{d(z, \text{supp}(\mu_N))} \quad \text{and} \quad \|\tilde{\mathbf{T}}_N(z)\| \leq \frac{\tilde{C}}{d(z, \text{supp}(\mu_N) \cup \{0\})}, \quad (12)$$

where $d(z, \mathcal{E})$ is the distance of z to a set \mathcal{E} , and C, \tilde{C} are two positive constants independent of N, z . We also have, for all $z \in \mathbb{C}$,

$$|1 + \sigma^2 c_N m_N(z)|^{-1} \leq 2, \quad (13)$$

and

$$\min_{k=1, \dots, M} |\lambda_{k,N} - w_N(z)| \geq \frac{d(z, \text{supp}(\mu_N))}{2}. \quad (14)$$

Note finally the two useful identities

$$1 + \sigma^2 c_N m_N(z) = \frac{1}{1 - \sigma^2 c_N f_N(w_N(z))} \text{ and } 1 + \sigma^2 c_N \tilde{m}_N(z) = \frac{1}{1 - \sigma^2 c_N \tilde{f}_N(w_N(z))}, \quad (15)$$

where f_N is defined by (7) and $\tilde{f}_N(w) = \frac{1}{M} \text{tr} (\mathbf{B}_N^* \mathbf{B}_N - w \mathbf{I})^{-1}$.

To conclude this section, we introduce some quantities which will appear during the computations of the CLT. We define

$$u_N(z_1, z_2) = \frac{\sigma^2}{N} \text{tr} \frac{\mathbf{T}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{T}_N(z_2)}{(1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2))} \quad (16)$$

as well as

$$v_N(z_1, z_2) = \frac{\sigma^2}{N} \text{tr} \mathbf{T}_N(z_1) \mathbf{T}_N(z_2) \quad \text{and} \quad \tilde{v}_N(z_1, z_2) = \frac{\sigma^2}{N} \text{tr} \tilde{\mathbf{T}}_N(z_1) \tilde{\mathbf{T}}_N(z_2). \quad (17)$$

Finally, we define

$$\Delta_N(z_1, z_2) = (1 - u_N(z_1, z_2))^2 - z_1 z_2 v_N(z_1, z_2) \tilde{v}_N(z_1, z_2). \quad (18)$$

The last quantity $\Delta_N(z_1, z_2)$ satisfies moreover the following bounds.

Lemma 1. *For all $z_1, z_2 \notin \text{supp}(\mu_N)$ such that $z_1 \neq z_2$, $\Delta_N(z_1, z_2)$ can be represented as*

$$\Delta_N(z_1, z_2) = \frac{z_1 - z_2}{w_N(z_1) - w_N(z_2)}. \quad (19)$$

Moreover, if there exists a closed set \mathcal{E} independent of N s.t. $\text{supp}(\mu_N) \subset \mathcal{E}$ for all large N , and if \mathcal{K} is a compact set s.t. $\mathcal{K} \subset \mathbb{C} \setminus (\{0\} \cup \mathcal{E})$, then

$$\limsup_{N \rightarrow \infty} \sup_{z_1, z_2 \in \mathcal{K}} |u_N(z_1, z_2)| < 1 \quad (20)$$

and

$$0 < \liminf_{N \rightarrow \infty} \inf_{z_1, z_2 \in \mathcal{K}} |\Delta_N(z_1, z_2)| \leq \limsup_{N \rightarrow \infty} \sup_{z_1, z_2 \in \mathcal{K}} |\Delta_N(z_1, z_2)| < \infty. \quad (21)$$

Finally, we also have

$$\left| \frac{\Delta_N(z_1, z_2)}{(1 - u_N(z_1, z_2))^2} - 1 \right| < 1, \quad (22)$$

for all $z_1, z_2 \in \mathcal{K}$.

Lemma 1 is proved in appendix 5.2.

2.3 Separation of the sample eigenvalues

In this section, we review some existing results concerning the location of the sample eigenvalues.

The following terminology will be used in the remainder: an eigenvalue $\lambda_{k,N}$ of $\mathbf{B}_N \mathbf{B}_N^*$ is *associated with the interval* $[x_{q,N}^-, x_{q,N}^+]$ of the support of μ_N if $w_N(x_{q,N}^-) < \lambda_{k,N} < w_N(x_{q,N}^+)$. It turns out that the "noise eigenvalue" 0 is always associated with the first interval $[x_{1,N}^-, x_{1,N}^+]$ since $w_N(x_{1,N}^-) < 0 < w_N(x_{1,N}^+)$ (see (8)), which is thus called in this context "noise cluster". Moreover, each "signal eigenvalue" $\lambda_{1,N}, \dots, \lambda_{K,N}$ is associated with a unique interval $[x_{q,N}^-, x_{q,N}^+]$ for $q = 1, \dots, Q$; in particular, a signal eigenvalue may be associated with the "noise cluster" while two signal eigenvalues may be associated with the same interval. We now introduce the two following additional assumptions, which informally ensure that the K signal eigenvalues $\lambda_{1,N}, \dots, \lambda_{K,N}$ will not be associated with the noise cluster, that is, will be separated from the "noise eigenvalue" 0 ($\lambda_{K+1,N}, \dots, \lambda_{M,N}$) for large N . This assumption will be necessary to guaranty the consistency of the subspace estimator introduced in the forthcoming sections.

Assumption A-1: *There exists $t_1^-, t_1^+, t_2^-, t_2^+$ s.t.*

$$0 < t_1^- < \liminf_{N \rightarrow \infty} x_{1,N}^- < \limsup_{N \rightarrow \infty} x_{1,N}^+ < t_1^+ < t_2^- < \liminf_{N \rightarrow \infty} x_{2,N}^- < \limsup_{N \rightarrow \infty} x_{2,N}^+ < t_2^+.$$

Assumption **A-1** thus ensures that the noise cluster remains asymptotically separated from the other intervals in the support of μ_N , as $N \rightarrow \infty$. From (8) and the fact that w_N is increasing on $\mathbb{R} \setminus \mu_N$, we have $w_N(t_1^-) < 0 < w_N(t_1^+) < w_N(t_2^-)$ for all large N . The second assumption **A-2** is related to the behaviour of the signal eigenvalues.

Assumption A-2: *For all large N , 0 is the unique eigenvalue of $\mathbf{B}_N \mathbf{B}_N^*$ associated with the noise cluster, i.e.*

$$w_N(t_2^-) < \lambda_{K,N}.$$

Note that this assumption implies that $\liminf_{N \rightarrow \infty} \lambda_{K,N} > 0$, thus ensuring that noise and signal eigenvalues are asymptotically separated (see lemma 2 below). These separation conditions have a direct consequence on the localization of the eigenvalues of the matrix $\Sigma_N \Sigma_N^*$. Indeed, it was shown in [16] that under assumptions **A-1** and **A-2**,

$$\hat{\lambda}_{1,N}, \dots, \hat{\lambda}_{M-K,N} \in [t_1^-, t_1^+] \quad \text{and} \quad \hat{\lambda}_{M-K+1,N}, \dots, \hat{\lambda}_{M,N} \in [t_2^-, t_2^+], \quad (23)$$

with probability one, for N large, i.e. the "noise sample eigenvalues" split from the "signal sample eigenvalues". An illustration of the density of μ_N and the localization of the sample eigenvalues (23) is given in figure 2.

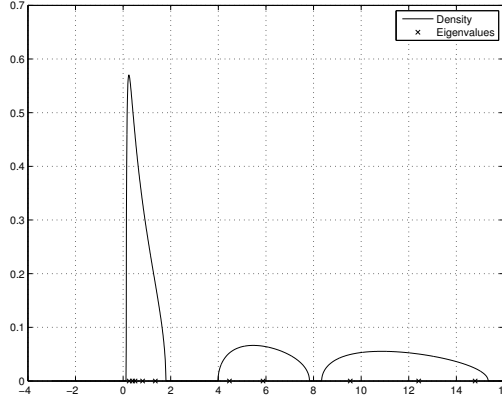


Figure 2: Density of μ_N and sample eigenvalues of $\Sigma_N \Sigma_N^*$ for one trial. The parameters are $M = 10$, $N = 20$, $\sigma = 1$ and the eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ are 0 (with multiplicity 5), 5 (with multiplicity 2) and 10 (with multiplicity 3)

Functions ϕ_N represents in some sense a link between the support of μ_N and the eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$. Figure 3 shows the consequence of assumption **A-2** on the behaviour of $\phi_N(w)$ near $w = 0$.

To conclude this part, we show that the separation conditions **A-1** and **A-2** ensure the effective separation between signal and noise eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$.

Lemma 2. *Assume the separation conditions **A-1** and **A-2** hold. Then,*

$$\liminf_{N \rightarrow \infty} \lambda_{K,N} > 0.$$

Proof. Assume the converse. Then there exists a subsequence $\varphi(N)$ such that $\lambda_{K,\varphi(N)} \rightarrow_N 0$. From the condition **A-2**, we have for all large N

$$w_N(t_1^-) < 0 < w_N(t_1^+) < w_N(t_2^-) < \lambda_{K,N},$$

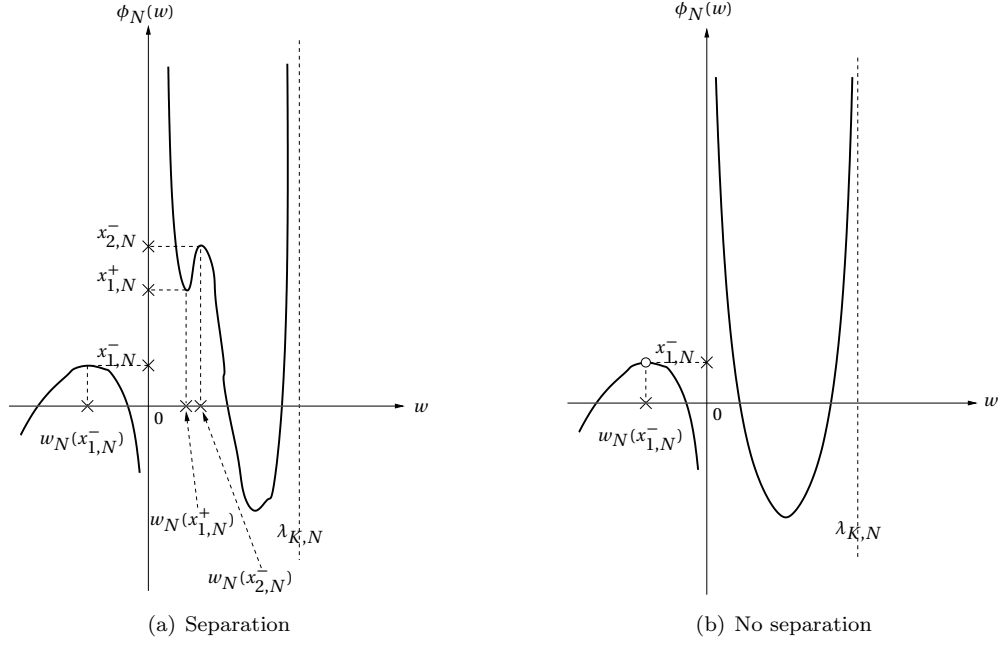


Figure 3: Typical example of the behaviour of ϕ_N near 0, when assumption **A-2** is satisfied (a), or not (b).

and the condition **A-1** ensures the existence of $x, y \in (t_1^+, t_2^-)$, with $x < y$, such that

$$w_{\varphi(N)}(y) - w_{\varphi(N)}(x) \xrightarrow[N \rightarrow \infty]{} 0. \quad (24)$$

But using (19) and (21) in lemma 1 contradicts (24). \square

2.4 The spiked model case: fixed rank

When K is constant with respect to N , the results of the previous sections can be simplified. Indeed, in this case, we have for all $z \in \mathbb{C} \setminus \mathbb{R}^+$,

$$m_N(z) - m(z) \xrightarrow[N \rightarrow \infty]{} 0, \quad (25)$$

where $m(z)$ satisfies the equation

$$m(z) = \frac{1}{-z(1 + \sigma^2 c m(z)) + \sigma^2(1 - c)}, \quad (26)$$

and is the Stieltjes transform of the Marchenko-Pastur distribution [10], with support $[\sigma^2(1 - \sqrt{c})^2, \sigma^2(1 + \sqrt{c})^2]$. An illustration of the Marchenko-Pastur distribution is given in figure 4.

For any compact $\mathcal{K} \in \mathbb{C} \setminus ([t_1^-, t_1^+] \cup [t_2^-, t_2^+])$, (25) can be strengthened with

$$\sup_{z \in \mathcal{K}} |m_N(z) - m(z)| \xrightarrow[N \rightarrow \infty]{} 0, \quad (27)$$

Simple algebra allows to rewrite the usual quantities in a simpler way. Indeed, we will have (in the same way as for (15))

$$1 + \sigma^2 c m(z) = \frac{w(z)}{w(z) + \sigma^2 c},$$

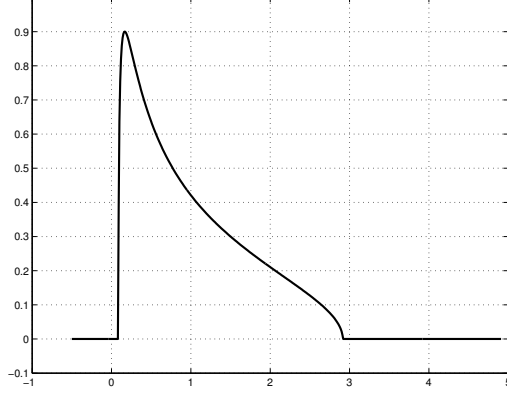


Figure 4: Density of the Marchenko-Pastur distribution in the case where $c = 0.5$, $\sigma = 1$.

with $w(z)$ given by

$$w(z) = z \left(1 + \sigma^2 c m(z)\right)^2 + \sigma^2 (1 - c) \left(1 + \sigma^2 c m(z)\right). \quad (28)$$

As for (5), equation (26) can be rewritten as

$$\phi(w(z)) = z, \quad (29)$$

where

$$\phi(w) = \frac{(w + \sigma^2 c)(w + \sigma^2)}{w}. \quad (30)$$

Of course, the boundary points of the support of the Marchenko-Pastur distribution, namely $\sigma^2(1 - \sqrt{c})^2$ and $\sigma^2(1 + \sqrt{c})^2$, coincides with the local extrema of ϕ , and with respective preimages

$$w(\sigma^2(1 - \sqrt{c})^2) = -\sigma^2 \sqrt{c} \text{ and } w(\sigma^2(1 + \sqrt{c})^2) = \sigma^2 \sqrt{c}.$$

The function w is continuous on \mathbb{R} , real and increasing on $\mathbb{R} \setminus [\sigma^2(1 - \sqrt{c})^2, \sigma^2(1 + \sqrt{c})^2]$. Figure (5) provides an illustration of the behaviour of the density of μ_N when K is fixed and $N \rightarrow \infty$. In the case when K is fixed, the separation assumptions **A-1** and **A-2** have a useful consequence on the behaviour of the smallest non zero eigenvalue of $\mathbf{B}_N \mathbf{B}_N^*$.

Lemma 3. *Assume K independent of N . Then, the separation condition **A-1** and **A-2** hold iff*

$$\liminf_{N \rightarrow \infty} \lambda_{K,N} > \sigma^2 \sqrt{c}. \quad (31)$$

The proof of lemma 3 is deferred in Appendix 5.3.

In the special situation where the non-zero eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ converge to some different limits, i.e.

$$\lambda_{k,N} \rightarrow_N \lambda_k > 0 \quad (32)$$

for all $k = 1, \dots, K$, with $\sigma^2 \sqrt{c} < \lambda_K < \dots < \lambda_1$, it is shown in [9] that the number Q of clusters in the support of μ_N is exactly $K + 1$ for N large and in this case, the "noise" eigenvalue 0 is the unique eigenvalue associated with the "noise" cluster $[x_{1,N}^-, x_{1,N}^+]$. Therefore, assumptions **A-1** and **A-2** are ensured in this case. It is also proved that $x_{k+1,N}^\pm \rightarrow \phi(\lambda_k)$, and using a refinement of (23), [9] also showed that (see also Benaych & Nadakuditi [2]) that the K largest sample eigenvalues $\hat{\lambda}_{1,N}, \dots, \hat{\lambda}_{K,N}$ split from the $M - K$ smallest eigenvalues $\hat{\lambda}_{K+1,N}, \dots, \hat{\lambda}_{M,N}$ and

$$\hat{\lambda}_{k,N} \xrightarrow[N \rightarrow \infty]{a.s.} \phi(\lambda_k), \quad (33)$$

while $\hat{\lambda}_{K+1,N} \rightarrow_N \sigma^2(1 + \sqrt{c})^2$ and $\hat{\lambda}_{M,N} \rightarrow_N \sigma^2(1 - \sqrt{c})^2$ a.s.

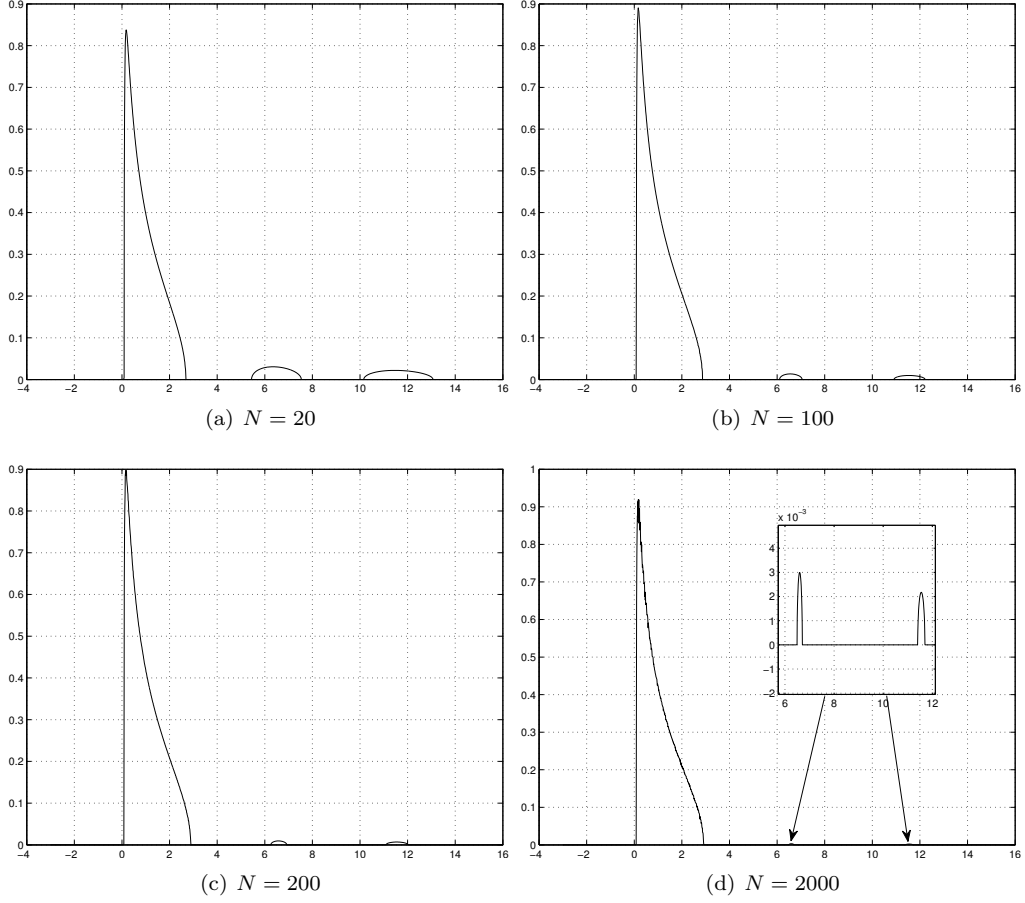


Figure 5: Density of μ_N when $K = 2$, $c_N = 0.5$. The two non zero eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ are 5 and 10.

Finally, we notice that $\Delta_N(z_1, z_2)$ defined in (18) will satisfy, when K is constant

$$\sup_{z_1, z_2 \in \mathcal{K}} |\Delta_N(z_1, z_2) - \Delta(z_1, z_2)| \xrightarrow{N \rightarrow \infty} 0,$$

where $\Delta(z_1, z_2)$ is given by

$$\Delta(z_1, z_2) = 1 - \frac{\sigma^4 c}{w(z_1)w(z_2)}. \quad (34)$$

The properties given in lemma 1 are of course valid for $\Delta(z_1, z_2)$: in particular, we have

$$\Delta(z_1, z_2) = \frac{z_1 - z_2}{w(z_1) - w(z_2)},$$

for all $z_1 \neq z_2$ and $z_1, z_2 \notin [\sigma^2(1 - \sqrt{c})^2, \sigma^2(1 + \sqrt{c})^2]$, as well as

$$0 < \inf_{z_1, z_2 \in \mathcal{K}} |\Delta(z_1, z_2)| \leq \sup_{z_1, z_2 \in \mathcal{K}} |\Delta(z_1, z_2)| < \infty$$

and

$$|\Delta(z_1, z_2) - 1| < 1,$$

for all $z_1, z_2 \in \mathcal{K}$ with \mathcal{K} a compact set such that $\mathcal{K} \subset \mathbb{C} \setminus (\{0\} \cup [\sigma^2(1 - \sqrt{c})^2, \sigma^2(1 + \sqrt{c})^2])$.

2.5 Contour integrals

Throughout the paper, we will deal with integrals of the form

$$I_N = \frac{1}{2\pi i} \int_{\partial\mathcal{R}} \Psi_N(w_N(z)) w'_N(z) dz \quad (35)$$

where $\partial\mathcal{R}$ is the clockwise oriented boundary of a rectangle \mathcal{R} intersecting the real axis at two points $t_2^- - \epsilon$, $t_2^+ + \epsilon$ with $\epsilon > 0$ such that $t_2^- > t_1^+ + \epsilon$, and where Ψ_N is a meromorphic function with poles contained in the set $\{\lambda_{1,N}, \dots, \lambda_{K,N}, 0\}$. It is shown in [16] from assumptions **A-1** and **A-2** that the set $w_N(\partial\mathcal{R})$ is a closed piecewise \mathcal{C}^1 path intersecting the real axis at points $w_N(t_2^- - \epsilon)$, $w_N(t_2^+ + \epsilon)$, enclosing the non-zero eigenvalues of $\mathbf{B}_N \mathbf{B}_N^*$ with winding number -1 and leaving 0 outside, for N large. Therefore, for all large N , a change of variable and residue theorem lead to

$$I_N = \frac{1}{2\pi i} \oint_{w_N(\partial\mathcal{R})} \Psi_N(w) dw = - \sum_{k=1}^K \text{Res}(\Psi_N, \lambda_{k,N}), \quad (36)$$

where $\text{Res}(\Psi_N, \lambda)$ is the residue of Ψ_N at λ . Note that in the case of the spiked models (see section 2.4) where K is fixed with respect to N , and under assumptions **A-1** and **A-2**, the previous result still holds by replacing $w_N(z)$ with $w(z)$ and $w'_N(z)$ with $w'(z)$.

3 Noise subspace estimation

In this section, we review the results of [16], [6] on the consistent subspace estimation, in the asymptotic regime where the number of antennas $M = M(N)$ is a function of the number of samples N such that $c_N = \frac{M}{N} \rightarrow c \in (0, 1)$ as $N \rightarrow \infty$.

3.1 Consistent estimation

Noise subspace estimation consists in our case in estimating the quantity

$$\eta_N = \mathbf{d}_{1,N}^* \mathbf{\Pi}_N \mathbf{d}_{2,N} = \mathbf{d}_{1,N}^* \left(\mathbf{I} - \sum_{k=1}^K \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \right) \mathbf{d}_{2,N}, \quad (37)$$

where $(\mathbf{d}_{1,N})$, $(\mathbf{d}_{2,N})$ are two sequences of deterministic vectors such that $\sup_N \|\mathbf{d}_{1,N}\|, \sup_N \|\mathbf{d}_{2,N}\| < \infty$.

We recall that the traditional estimator based on the SCM $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}$ is defined by

$$\hat{\eta}_N^{(t)} = \mathbf{d}_{1,N}^* \left(\mathbf{I} - \sum_{k=1}^K \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \right) \mathbf{d}_{2,N}. \quad (38)$$

It was shown that under the separation assumptions **A-1** and **A-2**, the quantity (37) can be written in terms of the following integral

$$\eta_N = \mathbf{d}_{1,N}^* \left(\mathbf{I} - \frac{1}{2\pi i} \oint_{\partial\mathcal{R}} \mathbf{T}_N(z) \frac{w'_N(z)}{1 + \sigma^2 c_N m_N(z)} dz \right) \mathbf{d}_{2,N}, \quad (39)$$

where $\partial\mathcal{R}$ is the clockwise oriented boundary of the rectangle

$$\mathcal{R} = \{x + iy : x \in [t_2^- - \epsilon, t_2^+ + \epsilon], y \in [-\delta, \delta]\}, \quad (40)$$

with $\epsilon > 0$ s.t. $t_1^+ + \epsilon < t_2^-$ and $\delta > 0$. By defining

$$\hat{w}_N(z) = z(1 + \sigma^2 c_N \hat{m}_N(z))^2 - \sigma^2(1 + \sigma^2 c_N \hat{m}_N(z)), \quad (41)$$

it is shown in [16] that

$$\sup_{z \in \partial\mathcal{R}} \left| \mathbf{d}_{1,N}^* \mathbf{Q}_N(z) \mathbf{d}_{2,N} \frac{\hat{w}'_N(z)}{1 + \sigma^2 c_N \hat{m}_N(z)} - \mathbf{d}_{1,N}^* \mathbf{T}_N(z) \mathbf{d}_{2,N} \frac{w'_N(z)}{1 + \sigma^2 c_N m_N(z)} \right| \xrightarrow[N \rightarrow \infty]{a.s.} 0. \quad (42)$$

This of course readily implies that

$$\hat{\eta}_N - \eta_N \xrightarrow[N \rightarrow \infty]{a.s.} 0, \quad (43)$$

where

$$\hat{\eta}_N = \mathbf{d}_{1,N}^* \left(\mathbf{I} - \frac{1}{2\pi i} \oint_{\partial \mathcal{R}} \mathbf{Q}_N(z) \frac{\hat{w}'_N(z)}{1 + \sigma^2 c_N \hat{m}_N(z)} dz \right) \mathbf{d}_{2,N}. \quad (44)$$

Thus $\hat{\eta}_N$ is a consistent estimator of (37).

Remark 1. The integrand in (44) is meromorphic with poles at $\hat{\lambda}_{1,N}, \dots, \hat{\lambda}_{M,N}$ as well as at the zeros of the function $z \mapsto 1 + \sigma^2 c_N \hat{m}_N(z)$, denoted $\hat{\omega}_{1,N}, \dots, \hat{\omega}_{M,N}$. It is shown in [16] that these zeros are the eigenvalues of the matrix $\hat{\Omega}_N = \hat{\Lambda}_N + \frac{\sigma^2 c_N}{M} \mathbf{1}\mathbf{1}^T$, where $\hat{\Lambda}_N = \text{diag}(\hat{\lambda}_{1,N}, \dots, \hat{\lambda}_{M,N})$, and follow a property similar to (23), i.e

$$\hat{\omega}_{1,N}, \dots, \hat{\omega}_{M-K,N} \in [t_1^-, t_1^+] \quad \text{and} \quad \hat{\omega}_{M-K+1,N}, \dots, \hat{\omega}_{M,N} \in [t_2^-, t_2^+], \quad (45)$$

with probability one, for N large enough. This ensures that the integral can be solved using residue theorem, and an explicit formula in terms of $\hat{\mathbf{u}}_{k,N}, \hat{\lambda}_{k,N}$ and $\hat{\omega}_{k,N}$ was provided in [16] for the improved subspace estimator (44).

Remark 2. Originally, the estimator derived in [16] was based on the representation

$$\eta_N = \frac{1}{2\pi i} \oint_{\partial \tilde{\mathcal{R}}} \mathbf{d}_{1,N}^* \mathbf{T}_N(z) \mathbf{d}_{2,N} \frac{w'_N(z)}{1 + \sigma^2 c_N m_N(z)} dz,$$

where $\partial \tilde{\mathcal{R}}$ is the clockwise oriented boundary of the rectangle

$$\tilde{\mathcal{R}} = \{x + iy : x \in [t_1^- - \epsilon, t_1^+ + \epsilon], y \in [-\delta, \delta]\},$$

enclosing the noise cluster (the contour (39) enclosing the signal cluster). In that case, (42) with \mathcal{R} replaced by $\tilde{\mathcal{R}}$ still holds, and

$$\hat{\eta}_N = \frac{1}{2\pi i} \oint_{\partial \tilde{\mathcal{R}}} \mathbf{d}_{1,N}^* \mathbf{Q}_N(z) \mathbf{d}_{2,N} \frac{\hat{w}'_N(z)}{1 + \sigma^2 c_N \hat{m}_N(z)} dz.$$

Therefore, the subspace estimators of [16] and (44) coincide. We choose to keep the representation (44) (with contour enclosing the signal cluster), which will be simpler to analyze in the following.

When K is constant, a simpler estimator of the localization function can be obtained [6] [16, Sec. C]. Indeed, since

$$\hat{\eta}_N = \mathbf{d}_{1,N}^* \left(\mathbf{I} - \frac{1}{2\pi i} \oint_{\partial \mathcal{R}} \mathbf{Q}_N(z) \frac{w'(z)}{1 + \sigma^2 c m(z)} dz \right) \mathbf{d}_{2,N} + o(1),$$

a straightforward application of residue theorem leads

$$\hat{\eta}_N = \hat{\eta}_N^{(s)} + o(1)$$

with probability one, where

$$\hat{\eta}_N^{(s)} = \mathbf{d}_{1,N}^* \left(\mathbf{I} - \sum_{k=1}^K h(\hat{\lambda}_{k,N}) \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \right) \mathbf{d}_{2,N}, \quad (46)$$

and where $h(z)$ is given by

$$h(z) = \frac{w'(z)}{1 + \sigma^2 c m(z)} = \frac{w(z)(w(z) + \sigma^2 c)}{w(z)^2 - \sigma^4 c}.$$

3.2 CLT

3.2.1 The main result

Before stating the main result, we need to introduce some new quantities. For $k, \ell \in \{1, \dots, M\}$, let

$$\vartheta_N(k, \ell) = \frac{\sigma^2}{2} \left(\frac{1}{2\pi i} \right)^2 \oint_{\partial\mathcal{R}} \oint_{\partial\mathcal{R}} \frac{\theta_N^{(k, \ell)}(z_1, z_2) w'_N(z_1) w'_N(z_2)}{(\lambda_{k, N} - w_N(z_1)) (\lambda_{\ell, N} - w_N(z_1)) (\lambda_{k, N} - w_N(z_2)) (\lambda_{\ell, N} - w_N(z_2)) \Delta_N(z_1, z_2)} dz_1 dz_2, \quad (47)$$

with

$$\begin{aligned} \theta_N^{(k, \ell)}(z_1, z_2) = & z_1 z_2 (1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2)) \tilde{v}_N(z_1, z_2) \\ & + \frac{\lambda_{k, N} \lambda_{\ell, N} v_N(z_1, z_2)}{(1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2))} + (\lambda_{k, N} + \lambda_{\ell, N}) (1 - u_N(z_1, z_2)). \end{aligned} \quad (48)$$

We define the 2×2 matrix $\mathbf{\Gamma}_N(k, \ell)$ by

$$\mathbf{\Gamma}_N(k, \ell) = \begin{bmatrix} \operatorname{Re} \left(\eta_{k, N}^{(1, 2)} \eta_{\ell, N}^{(1, 2)} \right) + \frac{1}{2} \left(\eta_{k, N}^{(1, 1)} \eta_{\ell, N}^{(2, 2)} + \eta_{\ell, N}^{(1, 1)} \eta_{k, N}^{(2, 2)} \right) & -\operatorname{Im} \left(\eta_{k, N}^{(1, 2)} \eta_{\ell, N}^{(1, 2)} \right) \\ -\operatorname{Im} \left(\eta_{k, N}^{(1, 2)} \eta_{\ell, N}^{(1, 2)} \right) & -\operatorname{Re} \left(\eta_{k, N}^{(1, 2)} \eta_{\ell, N}^{(1, 2)} \right) + \frac{1}{2} \left(\eta_{k, N}^{(1, 1)} \eta_{\ell, N}^{(2, 2)} + \eta_{\ell, N}^{(1, 1)} \eta_{k, N}^{(2, 2)} \right) \end{bmatrix},$$

where $\eta_{k, N}^{(i, j)} = \mathbf{d}_{i, N}^* \mathbf{u}_{k, N} \mathbf{u}_{k, N}^* \mathbf{d}_{j, N}$, and we finally set

$$\mathbf{\Gamma}_N = \sum_{k=1}^M \sum_{\ell=1}^M \vartheta_N(k, \ell) \mathbf{\Gamma}_N(k, \ell). \quad (49)$$

The main result is the following.

Theorem 1. *Assume the separation conditions **A-1** and **A-2** hold. Then we have*

$$0 \leq \liminf_{N \rightarrow \infty} \min_{k, \ell} \vartheta_N(k, \ell) \leq \limsup_{N \rightarrow \infty} \max_{k, \ell} \vartheta_N(k, \ell) < \infty. \quad (50)$$

Moreover, if $0 < \liminf_N \frac{K}{N} < \limsup_N \frac{K}{N} < c$, then

$$\liminf_{N \rightarrow \infty} \min_{k, \ell} \vartheta_N(k, \ell) > 0, \quad (51)$$

and if K is independent of N , then

$$\begin{aligned} \vartheta_N(k, \ell) = & \frac{\sigma^4 c_N (\lambda_{k, N} \lambda_{\ell, N} + (\lambda_{k, N} + \lambda_{\ell, N}) \sigma^2 + \sigma^4) (\lambda_{k, N} \lambda_{\ell, N} + \sigma^4 c_N)}{2 (\lambda_{k, N}^2 - \sigma^4 c_N) (\lambda_{\ell, N}^2 - \sigma^4 c_N) (\lambda_{k, N} \lambda_{\ell, N} - \sigma^4 c_N)} (1 - \mathbb{1}_{[K+1, M]}(k) \mathbb{1}_{[K+1, M]}(\ell)) + \epsilon_N(k, \ell), \end{aligned} \quad (52)$$

with $\max_{k, \ell} |\epsilon_N(k, \ell)| \rightarrow_N 0$. Finally, let (ξ_N) be a deterministic bounded sequence and denote $\xi_N = [\operatorname{Re}(\xi_N), \operatorname{Im}(\xi_N)]^T$. Then,

$$\operatorname{Re}(\xi_N (\hat{\eta}_N - \eta_N)) = \mathcal{O}_{\mathbb{P}} \left(\sqrt{\frac{\xi_N^T \mathbf{\Gamma}_N \xi_N}{N}} \right) + o_{\mathbb{P}} \left(\frac{1}{\sqrt{N}} \right), \quad (53)$$

and if $\liminf_N \xi_N^T \mathbf{\Gamma}_N \xi_N > 0$, it holds that

$$\sqrt{N} \frac{\operatorname{Re}(\xi_N (\hat{\eta}_N - \eta_N))}{\sqrt{\xi_N^T \mathbf{\Gamma}_N \xi_N}} \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}_{\mathbb{R}}(0, 1). \quad (54)$$

The proof of theorem (1) is deferred to section 4.

3.2.2 Discussions and numerical examples

In this section, we discuss the consequences of theorem 1 and provide numerical examples illustrating the results.

We first remark that in the statement of theorem (1), the purpose of the constraint $\liminf_N \boldsymbol{\xi}_N^T \boldsymbol{\Gamma}_N \boldsymbol{\xi}_N > 0$ is to ensure that the fluctuations of $\hat{\eta}_N - \eta_N$ are $\mathcal{O}(N^{-1/2})$. Indeed, there exist several situations where the fluctuations can be faster than $\mathcal{O}(N^{-1/2})$. For example, in the case where K is independent of N and $\mathbf{d}_{1,N} = \mathbf{d}_{2,N} = \mathbf{u}_{M,N}$, then we see from (52) that $\mathbf{e}_1^T \boldsymbol{\Gamma}_N \mathbf{e}_1 = o(1)$ and thus

$$\operatorname{Re}(\hat{\eta}_N - \eta_N) = o_{\mathbb{P}}\left(\frac{1}{\sqrt{N}}\right).$$

The result of theorem (1) can be rephrased in a more precise way, by considering the fluctuations of the random vector $[\operatorname{Re}(\hat{\eta}_N - \eta_N), \operatorname{Im}(\hat{\eta}_N - \eta_N)]^T$. However, in this case, we have to take into account the possible "degenerate" situations, when the covariance matrix $\boldsymbol{\Gamma}_N$ is asymptotically singular. Since it is difficult to state general results in this case, the next corollary focuses on one important special case where $\mathbf{d}_{1,N} = \mathbf{d}_{2,N}$, i.e. the case of quadratic forms.

Corollary 1 (Quadratic forms). *Assume that $\mathbf{d}_{1,N} = \mathbf{d}_{2,N} = \mathbf{d}_N$, where (\mathbf{d}_N) is a sequence of deterministic vectors such that $\limsup_N \|\mathbf{d}_N\| < \infty$. Under the assumptions of theorem 1,*

$$\frac{\sqrt{N}(\hat{\eta}_N - \eta_N)}{\sqrt{2 \sum_{k=1}^M \sum_{\ell=1}^M \vartheta_N(k, \ell) \mathbf{d}_N^* \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \mathbf{d}_N \mathbf{d}_N^* \mathbf{u}_{\ell,N} \mathbf{u}_{\ell,N}^* \mathbf{d}_N}} \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}_{\mathbb{R}}(0, 1).$$

The result of corollary 1 is illustrated in figure 6 by comparing the empirical distribution of the quadratic form $\sqrt{N}(\hat{\eta}_N - \eta_N)$ (10^5 trials), where $\mathbf{d}_{1,N} = \mathbf{d}_{2,N} = \mathbf{e}_M$, with the normal distribution $\mathcal{N}_{\mathbb{R}}(0, 2\vartheta_N(M, M))$. The parameters are $M = 20$, $N = 40$, $\sigma = 1$ and the matrix $\mathbf{B}_N \mathbf{B}_N^*$ is diagonal with non-zero eigenvalues at 5 and 6.

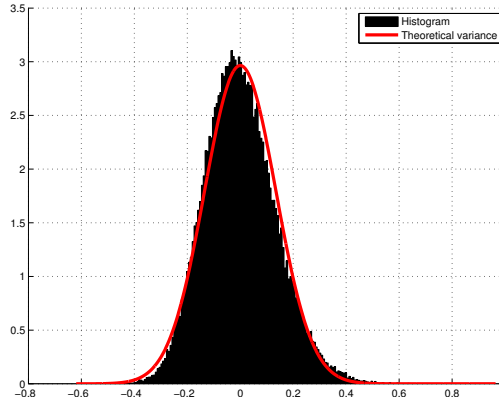


Figure 6: Empirical distribution of $\sqrt{N}(\hat{\eta}_N - \eta_N)$ (quadratic form)

In the "non-degenerate" case, we have to ensure that $\boldsymbol{\Gamma}_N$ is asymptotically non-singular. By computing the smallest eigenvalue of $\boldsymbol{\Gamma}_N$, this is equivalent to

$$\liminf_{N \rightarrow \infty} \left(\sum_{k, \ell} \left(\vartheta_N(k, \ell) \eta_{k,N}^{(1,1)} \eta_{\ell,N}^{(2,2)} \right) - \left| \sum_{k, \ell} \vartheta_N(k, \ell) \eta_{k,N}^{(1,2)} \eta_{\ell,N}^{(1,2)} \right| \right) > 0, \quad (55)$$

We therefore have the following result, by using the fact that

$$\operatorname{Re}(\boldsymbol{\xi}_N(\hat{\eta}_N - \eta_N)) = [\operatorname{Re}(\boldsymbol{\xi}_N) \quad \operatorname{Im}(\boldsymbol{\xi}_N)] \begin{bmatrix} \operatorname{Re}(\hat{\eta}_N - \eta_N) \\ -\operatorname{Im}(\hat{\eta}_N - \eta_N) \end{bmatrix} \quad (56)$$

Corollary 2 (Non-degenerate case). *Under the assumptions of theorem 1 and if (55) holds, then*

$$\sqrt{N} \mathbf{\Gamma}_N^{-1/2} \begin{bmatrix} \operatorname{Re}(\hat{\eta}_N - \eta_N) \\ -\operatorname{Im}(\hat{\eta}_N - \eta_N) \end{bmatrix} \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}_{\mathbb{R}^2}(\mathbf{0}, \mathbf{I}).$$

An illustration of corollary (2) is given in figure 7 where we have compared the empirical distribution of the bilinear form $\sqrt{N}\operatorname{Re}(\hat{\eta}_N - \eta_N)$ (10^5 trials), where $\mathbf{d}_{1,N} = \mathbf{e}_M$ and $\mathbf{d}_{2,N} = \mathbf{e}_{M-1}$, with $\mathcal{N}_{\mathbb{R}}(0, \mathbf{e}_1^T \mathbf{\Gamma}_N \mathbf{e}_1)$. The parameters are $M = 20$, $N = 40$, $\sigma = 1$ and the matrix $\mathbf{B}_N \mathbf{B}_N^*$ is diagonal with non-zero eigenvalues at 5 and 6 ($K = 2$).

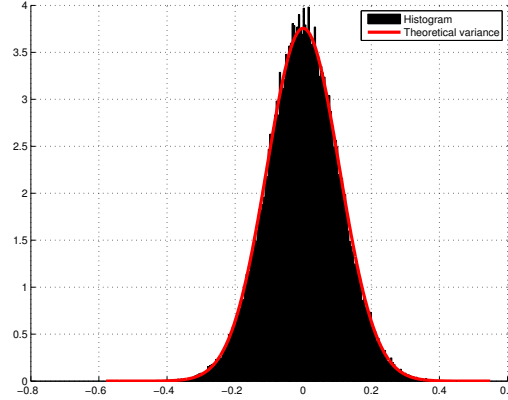


Figure 7: Empirical distribution of $\sqrt{N}\operatorname{Re}(\hat{\eta}_N - \eta_N)$ (bilinear form).

3.2.3 CLT for the traditional noise subspace estimate

To conclude section 3.2, we provide a CLT for the traditional noise subspace estimate, defined in (38) by

$$\hat{\eta}_N^{(t)} = \mathbf{d}_{1,N}^* \left(\mathbf{I} - \sum_{k=1}^K \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \right) \mathbf{d}_{2,N}.$$

From (23), almost surely for N large enough, the K largest eigenvalues $\hat{\lambda}_{1,N}, \dots, \hat{\lambda}_{K,N}$ are located inside the rectangle \mathcal{R} defined in (40), while the smallest $M - K$ remain outside \mathcal{R} . Therefore, for N large enough, almost surely,

$$\hat{\eta}_N^{(t)} = \mathbf{d}_{1,N}^* \left(\mathbf{I} - \frac{1}{2\pi i} \oint_{\partial \mathcal{R}} \mathbf{Q}_N(z) dz \right) \mathbf{d}_{2,N}.$$

As for (42), we have

$$\sup_{z \in \partial \mathcal{R}} |\mathbf{d}_{1,N}^* (\mathbf{Q}_N(z) - \mathbf{T}_N(z)) \mathbf{d}_{2,N}| \xrightarrow[N \rightarrow \infty]{a.s.} 0.$$

which immediately implies that

$$\hat{\eta}_N^{(t)} - \eta_N^{(t)} \xrightarrow[N \rightarrow \infty]{a.s.} 0,$$

where

$$\eta_N^{(t)} = \frac{1}{2\pi i} \oint_{\partial \mathcal{R}} \mathbf{d}_{1,N}^* \mathbf{T}_N(z) \mathbf{d}_{2,N} dz.$$

Define, as for (47),

$$\begin{aligned} \vartheta_N^{(t)}(k, \ell) = & \left(\frac{1}{2\pi i} \right)^2 \oint_{\partial \mathcal{R}} \oint_{\partial \mathcal{R}} \frac{\frac{\sigma^2}{2} \theta_N^{(k, \ell)}(z_1, z_2) (1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2))}{(\lambda_{k,N} - w_N(z_1)) (\lambda_{\ell,N} - w_N(z_1)) (\lambda_{k,N} - w_N(z_2)) (\lambda_{\ell,N} - w_N(z_2)) \Delta_N(z_1, z_2)} dz_1 dz_2, \end{aligned} \quad (57)$$

with $\Delta_N(z_1, z_2)$ and $\theta_N^{(k, \ell)}(z_1, z_2)$ defined respectively by (18) and (48). We define also

$$\mathbf{\Gamma}_N^{(t)} = \sum_{k=1}^M \sum_{\ell=1}^M \vartheta_N^{(t)}(k, \ell) \mathbf{\Gamma}_N(k, \ell), \quad (58)$$

as for (49). Then we have the following result :

Theorem 2. *Assume the separation conditions **A-1** and **A-2** hold, and that K is fixed with respect to N . Then $\max_{k, \ell \geq K+1} \vartheta_N^{(t)}(k, \ell) \rightarrow_N 0$,*

$$\vartheta_N^{(t)}(k, \ell) = \frac{\sigma^2 (\lambda_{k,N} + \sigma^2) (\lambda_{k,N}^2 - \sigma^4 c_N)}{2\lambda_{k,N}^2 (\lambda_{k,N} + \sigma^2 c_N)^2} + \epsilon_N(k, \ell)$$

for $k \leq K, \ell \geq K+1$, with $\max_{k \leq K, \ell \geq K+1} |\epsilon_N(k, \ell)| \rightarrow_N 0$, and

$$\vartheta_N^{(t)}(k, \ell) = \frac{\sigma^4 c_N \chi_N^{(t)}(k, \ell)}{2\lambda_{k,N} \lambda_{\ell,N} (\lambda_{k,N} + \sigma^2 c_N)^2 (\lambda_{\ell,N} + \sigma^2 c_N)^2 (\lambda_{k,N} \lambda_{\ell,N} - \sigma^4 c_N)} + o(1)$$

for $1 \leq k, \ell \leq K$, where $\chi_N^{(t)}(k, \ell)$ is defined by

$$\begin{aligned} \chi_N^{(t)}(k, \ell) = & \lambda_{k,N} \lambda_{\ell,N} (\lambda_{k,N} \lambda_{\ell,N} + \sigma^2 (\lambda_{k,N} + \lambda_{\ell,N}) + \sigma^4) ((1 + c_N)(\lambda_{k,N} \lambda_{\ell,N} + \sigma^4 c_N) + 2\sigma^2 c_N (\lambda_{k,N} + \lambda_{\ell,N})) \\ & - c (\lambda_{k,N} \lambda_{\ell,N} - \sigma^4 c_N) (\lambda_{k,N} \lambda_{\ell,N} + \sigma^2 (\lambda_{k,N} + \lambda_{\ell,N}) + \sigma^4 c_N)^2. \end{aligned}$$

Finally, let (ξ_N) be a deterministic bounded sequence and denote $\boldsymbol{\xi}_N = [\text{Re}(\xi_N), \text{Im}(\xi_N)]^T$. Then, if $\liminf_N \boldsymbol{\xi}_N^T \mathbf{\Gamma}_N^{(t)} \boldsymbol{\xi}_N > 0$, it holds that

$$\sqrt{N} \frac{\text{Re} \left(\xi_N \left(\hat{\eta}_N^{(t)} - \eta_N^{(t)} \right) \right)}{\sqrt{\boldsymbol{\xi}_N^T \mathbf{\Gamma}_N^{(t)} \boldsymbol{\xi}_N}} \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}_{\mathbb{R}}(0, 1).$$

The proof of Theorem 2, which follows step by step the proof of Theorem 1, is omitted.

4 Proof of theorem 1

This section is dedicated to prove theorem 1. Several long computations will be deferred to the appendix.

4.1 Regularization and confinement of the eigenvalues

To prove theorem 1, we will use the usual Levy's theorem and prove the convergence of the characteristic function. Since the moments of $\hat{\eta}_N$ may not be defined, due to the poles in the integrand of (44) (see remark 1), we first use a trick from [7], to force these poles to be away from the contour, and which does not modify the asymptotic distribution of $\hat{\eta}_N$.

Let $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}, [0, 1])$ s.t.

$$\varphi(\lambda) = \begin{cases} 1 & \text{for } \lambda \in [t_1^- - \frac{\epsilon}{3}, t_1^+ + \frac{\epsilon}{3}] \cup [t_2^- - \frac{\epsilon}{3}, t_2^+ + \frac{\epsilon}{3}] \\ 0 & \text{for } \lambda \in \mathbb{R} \setminus ([t_1^- - \frac{2\epsilon}{3}, t_1^+ + \frac{2\epsilon}{3}] \cup [t_2^- - \frac{2\epsilon}{3}, t_2^+ + \frac{2\epsilon}{3}]), \end{cases} \quad (59)$$

where ϵ is given in (40), and define the regularization coefficient

$$\chi_N = \det \varphi(\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^*) \det \varphi(\hat{\boldsymbol{\Omega}}_N) \quad (60)$$

(see remark 4 for the definition of $\hat{\Omega}_N$). From (23) and (45), we have $\chi_N = 1$ w.p.1 for N large and thus, for all $p \in \mathbb{N}$, we get

$$\hat{\eta}_N \chi_N^2 = \hat{\eta}_N + \mathcal{O}_{\mathbb{P}} \left(\frac{1}{N^p} \right).$$

Therefore, to obtain a CLT for $\hat{\eta}_N$, we only need to prove a CLT for $\hat{\eta}_N \chi_N^2$. Moreover, it is proved in [7] that

$$\sup_{z \in \partial \mathcal{R}} \mathbb{E} \left[\left| \mathbf{d}_{1,N}^* (\mathbf{Q}_N(z) - \mathbf{T}_N(z)) \mathbf{d}_{2,N} \right|^2 \chi_N \right] = \mathcal{O} \left(\frac{1}{N} \right), \quad (61)$$

$$\sup_{z \in \partial \mathcal{R}} \mathbb{E} \left[\left| \frac{\hat{w}'_N(z)}{1 + \sigma^2 c_N \hat{m}_N(z)} - \frac{w'_N(z)}{1 + \sigma^2 c_N m_N(z)} \right|^2 \chi_N \right] = \mathcal{O} \left(\frac{1}{N^2} \right). \quad (62)$$

Since $\frac{\hat{w}'_N(z)}{1 + \sigma^2 c_N \hat{m}_N(z)} \chi_N$ fluctuates less than the quadratic form $\mathbf{d}_{1,N}^* \mathbf{Q}_N(z) \mathbf{d}_{2,N} \chi_N$, we can replace it with $\frac{w'_N(z)}{1 + \sigma^2 c_N m_N(z)}$ without modifying any asymptotic second order results. Indeed, it is easy to see from (62) that

$$\hat{\eta}_N \chi_N^2 = \mathbf{d}_{1,N}^* \left(\mathbf{I} - \frac{1}{2\pi i} \oint_{\partial \mathcal{R}} \mathbf{Q}_N(z) \chi_N \frac{w'_N(z)}{1 + \sigma^2 c_N m_N(z)} dz \right) \mathbf{d}_{2,N} + \mathcal{O}_{\mathbb{P}} \left(\frac{1}{N} \right),$$

and the problem reduces finally to obtain the asymptotic distribution of

$$\hat{\gamma}_N = \frac{1}{2\pi i} \oint_{\partial \mathcal{R}} \mathbf{d}_{1,N}^* \mathbf{Q}_N(z) \mathbf{d}_{2,N} \chi_N \frac{w'_N(z)}{1 + \sigma^2 c_N m_N(z)} dz. \quad (63)$$

Remark 3. If $\sqrt{N}(c_N - c) \rightarrow 0$, it can be proved that

$$\sup_{z \in \mathcal{R}} \left| \frac{w'_N(z)}{1 + \sigma^2 c_N m_N(z)} - \frac{w'(z)}{1 + \sigma^2 c m(z)} \right| = o \left(\frac{1}{\sqrt{N}} \right),$$

and the simplified estimator (46) $\hat{\eta}_N^{(s)}$ derived in [6] will have the same asymptotic fluctuations as $\hat{\eta}_N$.

In the remainder, we denote by $\psi_N(u)$ the characteristic function defined on \mathbb{R} by

$$\psi_N(u) = \mathbb{E} \left[\exp \left(iu \sqrt{N} \operatorname{Re} (\xi_N \hat{\gamma}_N) \right) \right],$$

where (ξ_N) is a deterministic sequence such that $\limsup_N |\xi_N| < \infty$.

Finally, we recall two useful properties from [7, Prop. 3.3]:

$$\mathbb{E} [\hat{\gamma}_N - (\mathbf{d}_{1,N}^* \mathbf{d}_{2,N} - \eta_N)] = \mathcal{O} \left(\frac{1}{N^{3/2}} \right) \text{ and } \mathbb{E} |\hat{\gamma}_N - (\mathbf{d}_{1,N}^* \mathbf{d}_{2,N} - \eta_N)|^2 = \mathcal{O} \left(\frac{1}{N} \right). \quad (64)$$

4.2 The differential equation

We first prove that the characteristic function $\psi_N(u)$ satisfies the differential equation of a Gaussian characteristic function, up to an error term.

Remark 4. Note that in the expression of $\psi_N(u)$, we can assume for ease of reading and without loss of generality that $\xi_N = 1$ (by considering vectors $\mathbf{d}_{1,N} \overline{\xi_N}^{1/2}$ and $\mathbf{d}_{2,N} \xi_N^{1/2}$ in the bilinear form).

In the following, $\epsilon_N(u, z_1, z_2)$ will denote a complex generic continuous function defined on $\mathbb{R} \times \partial \mathcal{R} \times \partial \mathcal{R}$, such that $u \mapsto \epsilon_N(u, z_1, z_2)$ is continuously differentiable, and

$$\limsup_{N \rightarrow \infty} \sup_{(z_1, z_2) \in \partial \mathcal{R} \times \partial \mathcal{R}} \left\{ |\epsilon_N(u, z_1, z_2)|, \left| \frac{\partial \epsilon_N(u, z_1, z_2)}{\partial u} \right| \right\} < P(u),$$

with $P(u)$ a polynomial with positive coefficients. $\epsilon_N(u, z_1, z_2)$ may take different values from one line to another. We will also keep the notation $\epsilon_N(u)$, $\epsilon_N(z_1, z_2)$, $\epsilon_N(u, z_1)$ if $\epsilon_N(u, z_1, z_2)$ does not depend on (z_1, z_2) , u or z_2 .

Using dominated convergence and Fubini's theorem, the derivative $\psi'_N(u)$ writes

$$\psi'_N(u) = \frac{i\sqrt{N}}{2} \frac{1}{2\pi i} \oint_{\partial\mathcal{R}} \mathbb{E} \left[(\mathbf{d}_{1,N}^* \mathbf{Q}_N(z) \mathbf{d}_{2,N} + \mathbf{d}_{2,N}^* \mathbf{Q}_N(z) \mathbf{d}_{1,N}) \chi_N e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] \frac{w'_N(z)}{1 + \sigma^2 c_N m_N(z)} dz, \quad (65)$$

so that we need to develop the term $\mathbb{E} [\mathbf{d}_{1,N}^* \mathbf{Q}_N(z) \mathbf{d}_{2,N} \chi_N e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)}]$. By standard computations deferred to appendix 5.1.1, we obtain

$$\begin{aligned} \psi'_N(u) = & \left(i\sqrt{N}\text{Re}(\mathbf{d}_{1,N}^* \mathbf{d}_{2,N} - \eta_N) - \frac{u\sigma^2}{4} \left(\frac{1}{2\pi i} \right)^2 \oint_{\partial\mathcal{R}} \oint_{\partial\mathcal{R}} \frac{(\mu_N(z_1, z_2) + \tilde{\mu}_N(z_1, z_2)) w'_N(z_1) w'_N(z_2)}{(1 + \sigma^2 c_N m_N(z_1))^2 (1 + \sigma^2 c_N m_N(z_2))^2} dz_1 dz_2 \right) \psi_N(u) \\ & + \frac{\epsilon_N(u)}{\sqrt{N}}, \end{aligned} \quad (66)$$

where the quantity $\mu_N(z_1, z_2)$ is given by

$$\begin{aligned} \mu_N(z_1, z_2) = & \sum_{\substack{i,j=1 \\ i \neq j}}^2 \left[\mathbf{d}_{i,N}^* \mathbf{T}_N(z_1) \mathbf{T}_N(z_2) \mathbf{d}_{j,N} \mathbb{E} [\mathbf{d}_{i,N}^* \mathbf{Q}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{j,N} \chi_N] \right. \\ & + \mathbf{d}_{i,N}^* \mathbf{T}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{T}_N(z_2) \mathbf{d}_{j,N} \mathbb{E} [\mathbf{d}_{i,N}^* \mathbf{Q}_N(z_1) \mathbf{Q}_N(z_2) \mathbf{d}_{j,N} \chi_N] \\ & \left. + s_N(z_1, z_2) (1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2)) \mathbf{d}_{i,N}^* \mathbf{T}_N(z_1) \mathbf{T}_N(z_2) \mathbf{d}_{j,N} \mathbb{E} [\mathbf{d}_{i,N}^* \mathbf{Q}_N(z_1) \mathbf{Q}_N(z_2) \mathbf{d}_{j,N} \chi_N] \right], \end{aligned} \quad (67)$$

and $\tilde{\mu}_N(z_1, z_2)$ is given by

$$\begin{aligned} \tilde{\mu}_N(z_1, z_2) = & \sum_{\substack{i,j=1 \\ i \neq j}}^2 \left[\mathbf{d}_{i,N}^* \mathbf{T}_N(z_1) \mathbf{T}_N(z_2) \mathbf{d}_{i,N} \mathbb{E} [\mathbf{d}_{j,N}^* \mathbf{Q}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{j,N} \chi_N] \right. \\ & + \mathbf{d}_{i,N}^* \mathbf{T}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{T}_N(z_2) \mathbf{d}_{i,N} \mathbb{E} [\mathbf{d}_{j,N}^* \mathbf{Q}_N(z_1) \mathbf{Q}_N(z_2) \mathbf{d}_{j,N} \chi_N] \\ & \left. + s_N(z_1, z_2) (1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2)) \mathbf{d}_{i,N}^* \mathbf{T}_N(z_1) \mathbf{T}_N(z_2) \mathbf{d}_{i,N} \mathbb{E} [\mathbf{d}_{j,N}^* \mathbf{Q}_N(z_1) \mathbf{Q}_N(z_2) \mathbf{d}_{j,N} \chi_N] \right], \end{aligned} \quad (68)$$

where $s_N(z_1, z_2)$ defined as

$$s_N(z_1, z_2) = \frac{\sigma^2}{N} \text{tr} \frac{\mathbf{I} - \frac{\mathbf{B}_N^* \mathbf{T}_N(z_1) \mathbf{B}_N}{1 + \sigma^2 c_N m_N(z_1)} - \frac{\mathbf{B}_N^* \mathbf{T}_N(z_2) \mathbf{B}_N}{1 + \sigma^2 c_N m_N(z_2)}}{(1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2))}. \quad (69)$$

It now remains to approximate $\mathbb{E} [\mathbf{d}_{i,N}^* \mathbf{Q}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{j,N} \chi_N]$ and $\mathbb{E} [\mathbf{d}_{i,N}^* \mathbf{Q}_N(z_1) \mathbf{Q}_N(z_2) \mathbf{d}_{j,N} \chi_N]$, and we introduce for that purpose the following quantity

$$r_N(z_1, z_2) = \frac{\sigma^2}{N} \text{tr} \frac{\mathbf{T}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{T}_N(z_2) \mathbf{B}_N \mathbf{B}_N^*}{((1 + \sigma^2 c_N m_N(z_1))^2 ((1 + \sigma^2 c_N m_N(z_2))^2)}. \quad (70)$$

Recall moreover the definitions of $u_N(z_1, z_2)$, $v_N(z_1, z_2)$ and $\tilde{v}_N(z_1, z_2)$ given by (16) and (17).

Proposition 1. *For all $z_1, z_2 \in \partial\mathcal{R}$, we have $s_N(z_1, z_2) + r_N(z_1, z_2) = z_1 z_2 \tilde{v}_N(z_1, z_2)$,*

$$\begin{aligned} \mathbb{E} [\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N] = & \frac{1 - u_N(z_1, z_2)}{\Delta_N(z_1, z_2)} \mathbf{d}_{1,N}^* \mathbf{T}_N(z_1) \mathbf{T}_N(z_2) \mathbf{d}_{2,N} + \frac{v_N(z_1, z_2)}{\Delta_N(z_1, z_2)} \frac{\mathbf{d}_{1,N}^* \mathbf{T}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{T}_N(z_2) \mathbf{d}_{2,N}}{(1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2))} + \frac{\epsilon_N(z_1, z_2)}{N}, \end{aligned} \quad (71)$$

and

$$\begin{aligned} \mathbb{E} [\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N] = \\ \frac{(1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2)) (u_N(z_1, z_2) s_N(z_1, z_2) + r_N(z_1, z_2))}{\Delta_N(z_1, z_2)} \mathbf{d}_{1,N}^* \mathbf{T}_N(z_1) \mathbf{T}_N(z_2) \mathbf{d}_{2,N} \\ + \frac{1 - s_N(z_1, z_2) v_N(z_1, z_2) - u_N(z_1, z_2)}{\Delta_N(z_1, z_2)} \mathbf{d}_{1,N}^* \mathbf{T}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{T}_N(z_2) \mathbf{d}_{2,N} + \frac{\epsilon_N(z_1, z_2)}{N}. \end{aligned} \quad (72)$$

Note that the inverse of $\Delta_N(z_1, z_2)$ is well defined thanks to (21) in lemma 1. The proof of proposition 1 is given in appendix 5.1.2. Using the expression of $\mathbb{E} [\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N]$ and $\mathbb{E} [\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N]$ in proposition 1 and the fact that $s_N(z_1, z_2) + r_N(z_1, z_2) = z_1 z_2 \tilde{v}_N(z_1, z_2)$, we further obtain

$$\begin{aligned} \mu_N(z_1, z_2) = \\ \sum_{\substack{i,j=1 \\ i \neq j}}^2 \left(\frac{z_1 z_2 (1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2)) \tilde{v}_N(z_1, z_2)}{\Delta_N(z_1, z_2)} (\mathbf{d}_{i,N}^* \mathbf{T}_N(z_1) \mathbf{T}_N(z_2) \mathbf{d}_{j,N})^2 \right. \\ \left. + \frac{v_N(z_1, z_2) (\mathbf{d}_{i,N}^* \mathbf{T}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{T}_N(z_2) \mathbf{d}_{j,N})^2}{(1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2)) \Delta_N(z_1, z_2)} \right. \\ \left. + \frac{2(1 - u_N(z_1, z_2))}{\Delta_N(z_1, z_2)} \mathbf{d}_{i,N}^* \mathbf{T}_N(z_1) \mathbf{T}_N(z_2) \mathbf{d}_{j,N} \mathbf{d}_{i,N}^* \mathbf{T}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{T}_N(z_2) \mathbf{d}_{j,N} \right) + \frac{\epsilon_N(z_1, z_2)}{N}, \end{aligned}$$

and

$$\begin{aligned} \tilde{\mu}_N(z_1, z_2) = \\ \sum_{\substack{i,j=1 \\ i \neq j}}^2 \left(\frac{z_1 z_2 (1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2)) \tilde{v}_N(z_1, z_2)}{\Delta_N(z_1, z_2)} \mathbf{d}_{i,N}^* \mathbf{T}_N(z_1) \mathbf{T}_N(z_2) \mathbf{d}_{i,N} \mathbf{d}_{j,N}^* \mathbf{T}_N(z_1) \mathbf{T}_N(z_2) \mathbf{d}_{j,N} \right. \\ \left. + \frac{v_N(z_1, z_2) \mathbf{d}_{i,N}^* \mathbf{T}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{T}_N(z_2) \mathbf{d}_{i,N} \mathbf{d}_{j,N}^* \mathbf{T}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{T}_N(z_2) \mathbf{d}_{j,N}}{(1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2)) \Delta_N(z_1, z_2)} \right. \\ \left. + \frac{2(1 - u_N(z_1, z_2))}{\Delta_N(z_1, z_2)} \mathbf{d}_{i,N}^* \mathbf{T}_N(z_1) \mathbf{T}_N(z_2) \mathbf{d}_{i,N} \mathbf{d}_{j,N}^* \mathbf{T}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{T}_N(z_2) \mathbf{d}_{j,N} \right) + \frac{\epsilon_N(z_1, z_2)}{N}, \end{aligned}$$

Going back to (65) and introducing again the deterministic sequence (ξ_N) (see remark 4), we finally obtain

$$\begin{aligned} \psi'_N(u) = \\ \left(i\sqrt{N} \text{Re} (\xi_N (\mathbf{d}_{1,N}^* \mathbf{d}_{2,N} - \eta_N)) - u \sum_{k=1}^M \sum_{\ell=1}^M \vartheta_N(k, \ell) \left(\text{Re} \left(\xi_N^2 \eta_{k,N}^{(1,2)} \eta_{\ell,N}^{(1,2)} \right) + \frac{|\xi_N|^2}{2} \left(\eta_{k,N}^{(1,1)} \eta_{\ell,N}^{(2,2)} + \eta_{k,N}^{(2,2)} \eta_{\ell,N}^{(1,1)} \right) \right) \right) \psi_N(u) \\ + \frac{\epsilon_N(u)}{\sqrt{N}}, \end{aligned} \quad (73)$$

where we recall that $\eta_{k,N}^{(i,j)} = \mathbf{d}_{i,N}^* \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \mathbf{d}_{j,N}$ and where $\vartheta_N(k, \ell)$ is given by

$$\begin{aligned} \vartheta_N(k, \ell) = \\ \frac{\sigma^2}{2} \left(\frac{1}{2\pi i} \right)^2 \oint_{\partial \mathcal{R}} \oint_{\partial \mathcal{R}} \frac{\theta_N^{(k,\ell)}(z_1, z_2) w'_N(z_1) w'_N(z_2)}{(\lambda_{k,N} - w_N(z_1)) (\lambda_{\ell,N} - w_N(z_1)) (\lambda_{k,N} - w_N(z_2)) (\lambda_{\ell,N} - w_N(z_2)) \Delta_N(z_1, z_2)} dz_1 dz_2, \end{aligned} \quad (74)$$

with

$$\begin{aligned} \theta_N^{(k,\ell)}(z_1, z_2) = & z_1 z_2 (1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2)) \tilde{v}_N(z_1, z_2) \\ & + \frac{\lambda_{k,N} \lambda_{\ell,N} v_N(z_1, z_2)}{(1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2))} + (\lambda_{k,N} + \lambda_{\ell,N}) (1 - u_N(z_1, z_2)). \end{aligned} \quad (75)$$

We can check easily that $\vartheta_N(k, \ell) \in \mathbb{R}$. By letting $\boldsymbol{\xi}_N = [\text{Re}(\xi_N), \text{Im}(\xi_N)]^T$, we obtain

$$\psi'_N(u) = \left(i\sqrt{N} \text{Re}(\xi_N(\mathbf{d}_{1,N}^* \mathbf{d}_{2,N} - \eta_N)) - u \boldsymbol{\xi}_N^T \boldsymbol{\Gamma}_N \boldsymbol{\xi}_N \right) \psi_N(u) + \frac{\epsilon_N(u)}{\sqrt{N}}, \quad (76)$$

where

$$\boldsymbol{\Gamma}_N = \sum_{k=1}^M \sum_{\ell=1}^M \vartheta_N(k, \ell) \boldsymbol{\Gamma}_N(k, \ell).$$

with the 2×2 matrix $\boldsymbol{\Gamma}_N(k, \ell)$ given by

$$\boldsymbol{\Gamma}_N(k, \ell) = \begin{bmatrix} \text{Re}(\eta_{k,N}^{(1,2)} \eta_{\ell,N}^{(1,2)}) + \frac{1}{2} (\eta_{k,N}^{(1,1)} \eta_{\ell,N}^{(2,2)} + \eta_{k,N}^{(2,2)} \eta_{\ell,N}^{(1,1)}) & -\text{Im}(\eta_{k,N}^{(1,2)} \eta_{\ell,N}^{(1,2)}) \\ -\text{Im}(\eta_{k,N}^{(1,2)} \eta_{\ell,N}^{(1,2)}) & -\text{Re}(\eta_{k,N}^{(1,2)} \eta_{\ell,N}^{(1,2)}) + \frac{1}{2} (\eta_{k,N}^{(1,1)} \eta_{\ell,N}^{(2,2)} + \eta_{k,N}^{(2,2)} \eta_{\ell,N}^{(1,1)}) \end{bmatrix}. \quad (77)$$

From the trivial inequality $|\text{Re}(z_1 z_2)| \leq \frac{1}{2} (|z_1|^2 + |z_2|^2)$ for $z_1, z_2 \in \mathbb{C}$, we have

$$\left| \text{Re}(\xi_N^2 \eta_{k,N}^{(1,2)} \eta_{\ell,N}^{(1,2)}) \right| \leq \frac{|\xi_N|^2}{2} (\eta_{k,N}^{(1,1)} \eta_{\ell,N}^{(2,2)} + \eta_{k,N}^{(2,2)} \eta_{\ell,N}^{(1,1)}),$$

which implies

$$\boldsymbol{\xi}_N^T \boldsymbol{\Gamma}_N(k, \ell) \boldsymbol{\xi}_N = \text{Re}(\xi_N^2 \eta_{k,N}^{(1,2)} \eta_{\ell,N}^{(1,2)}) + \frac{|\xi_N|^2}{2} (\eta_{k,N}^{(1,1)} \eta_{\ell,N}^{(2,2)} + \eta_{k,N}^{(2,2)} \eta_{\ell,N}^{(1,1)}) \geq 0, \quad (78)$$

in other words that $\boldsymbol{\Gamma}_N(k, \ell)$ is non-negative definite.

4.3 Asymptotics of $\vartheta_N(k, \ell)$

The purpose of this section is to prove (50), (51) and (52) for the coefficients $\vartheta_N(k, \ell)$.

Using the bounds (12), (13) and the fact that $|m'_N(z)| \leq d(z, \text{supp}(\mu_N))^{-2}$, it is easily shown that $\theta_N^{(k,\ell)}(z_1, z_2)$, defined in (48), satisfies

$$\limsup_{N \rightarrow \infty} \max_{k, \ell} \sup_{z_1, z_2 \in \partial \mathcal{R}} \left| \theta_N^{(k,\ell)}(z_1, z_2) w'_N(z_1) w'_N(z_2) \right| < \infty.$$

Moreover, from (14) and (21), we also have

$$\limsup_{N \rightarrow \infty} \max_{k, \ell} \sup_{z_1, z_2 \in \partial \mathcal{R}} \left| \frac{\Delta_N(z_1, z_2)^{-1}}{(\lambda_{k,N} - w_N(z_1)) (\lambda_{\ell,N} - w_N(z_1)) (\lambda_{k,N} - w_N(z_2)) (\lambda_{\ell,N} - w_N(z_2))} \right| < \infty.$$

Therefore, these bounds readily imply

$$\limsup_{N \rightarrow \infty} \max_{k, \ell} \vartheta_N(k, \ell) < \infty. \quad (79)$$

We now express the integrand of $\vartheta_N(k, \ell)$ as a series of functions which are separable and symmetric in z_1, z_2 (i.e. a function $g(z_1, z_2)$ is symmetric separable if it can be written as $g(z_1, z_2) = \tilde{g}(z_1) \tilde{g}(z_2)$). Notice that,

except for $\Delta_N(z_1, z_2)$, all the functions appearing in the integrand in the definition of $\vartheta_N(k, l)$ are trivially sums of separable functions from their very definition. From lemma 1, we have

$$\left| \frac{\Delta_N(z_1, z_2)}{(1 - u_N(z_1, z_2))^2} - 1 \right| < 1,$$

and by writing

$$\Delta_N(z_1, z_2)^{-1} = \frac{1}{(1 - u_N(z_1, z_2))^2 \left(1 - \frac{z_1 z_2 v_N(z_1, z_2) \tilde{v}_N(z_1, z_2)}{(1 - u_N(z_1, z_2))^2} \right)},$$

we obtain

$$\Delta_N(z_1, z_2)^{-1} = \sum_{k \in \mathbb{N}} \frac{(z_1 z_2 v_N(z_1, z_2) \tilde{v}_N(z_1, z_2))^k}{(1 - u_N(z_1, z_2))^{2k+2}}. \quad (80)$$

Using this time the fact that $|u_N(z_1, z_2)| < 1$ (see lemma 1), we can further write

$$\Delta_N(z_1, z_2)^{-1} = \sum_{k \in \mathbb{N}} \sum_{l_1, \dots, l_{2k+2} \in \mathbb{N}} (z_1 z_2 v_N(z_1, z_2) \tilde{v}_N(z_1, z_2))^k u_N(z_1, z_2)^{l_1 + \dots + l_{2k+2}}. \quad (81)$$

Since the functions u_N , v_N and \tilde{v}_N are continuous on the compact set $\partial\mathcal{R} \times \partial\mathcal{R}$, the bound previously derived shows that the series of functions defining (81) is uniformly convergent on $\partial\mathcal{R} \times \partial\mathcal{R}$. Consequently, we can rewrite the coefficients $\vartheta_N(k, \ell)$, defined in (47), as

$$\begin{aligned} \vartheta_N(k, \ell) = & \sum_{n \in \mathbb{N}} \sum_{l_1, \dots, l_{2n+2} \in \mathbb{N}} \left(\frac{1}{2\pi i} \right)^2 \oint_{\partial\mathcal{R}} \oint_{\partial\mathcal{R}} \frac{\sigma^2 \theta_N^{(k, \ell)}(z_1, z_2) (z_1 z_2 v_N(z_1, z_2) \tilde{v}_N(z_1, z_2))^n u_N(z_1, z_2)^{l_1 + \dots + l_{2n+2}} w'_N(z_1) w'_N(z_2)}{2 (\lambda_{k, N} - w_N(z_1)) (\lambda_{\ell, N} - w_N(z_1)) (\lambda_{k, N} - w_N(z_2)) (\lambda_{\ell, N} - w_N(z_2))} dz_1 dz_2, \end{aligned} \quad (82)$$

where $\theta_N^{(k, \ell)}(z_1, z_2)$ is defined in (48). In other words, we have written $\vartheta_N(k, \ell)$ as a convergent series of integrals of symmetric separable functions. Consequently, $\vartheta_N(k, \ell)$ can be written as a series of squared single integrals, i.e. there exists a sequence of continuous functions $(g_N^{(p)})_{p \in \mathbb{N}}$ defined on $\partial\mathcal{R}$ such that

$$\vartheta_N(k, \ell) = \sum_{p \in \mathbb{N}} \left(\frac{1}{2\pi i} \oint_{\partial\mathcal{R}} g_N^{(p)}(z) dz \right)^2$$

implying that $\vartheta_N(k, \ell) \geq 0$. This proves (50).

To prove (51), we rely on the series expansion (82) introduced above. Using only one of the three terms in the definition of $\theta_N^{(k, \ell)}(z_1, z_2)$ (see (48)), and by only considering $n = 0$ in the sum of the series in (82), we obtain the following lower-bound

$$\begin{aligned} \vartheta_N(k, \ell) \geq & \left(\frac{1}{2\pi i} \right)^2 \oint_{\partial\mathcal{R}} \oint_{\partial\mathcal{R}} \frac{\sigma^2 \lambda_{k, N} \lambda_{\ell, N} v_N(z_1, z_2) (1 + \sigma^2 c_N m_N(z_1))^{-1} (1 + \sigma^2 c_N m_N(z_2))^{-1} w'_N(z_1) w'_N(z_2)}{2 (\lambda_{k, N} - w_N(z_1)) (\lambda_{\ell, N} - w_N(z_1)) (\lambda_{k, N} - w_N(z_2)) (\lambda_{\ell, N} - w_N(z_2))} dz_1 dz_2, \end{aligned} \quad (83)$$

From the definition of $v_N(z_1, z_2)$ (see (17)), we have

$$\frac{v_N(z_1, z_2)}{(1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2))} = \frac{\sigma^2}{N} \sum_{m=1}^M \frac{1}{(\lambda_{m, N} - w_N(z_1)) (\lambda_{m, N} - w_N(z_2))},$$

and a usual change of variable gives

$$\vartheta_N(k, \ell) \geq \lambda_{k, N} \lambda_{\ell, N} \frac{\sigma^4}{2N} \sum_{m=1}^M \left(\frac{1}{2\pi i} \oint_{w_N(\partial\mathcal{R})} \frac{dw}{(\lambda_{m, N} - w) (\lambda_{k, N} - w) (\lambda_{\ell, N} - w)} \right)^2.$$

Residue's theorem thus implies that for $k, \ell \leq K$,

$$\vartheta_N(k, \ell) \geq \frac{M - K}{N} \frac{\sigma^4}{2\lambda_{k,N}\lambda_{\ell,N}}. \quad (84)$$

In the same way, for $k, \ell \geq K + 1$, we obtain

$$\begin{aligned} \vartheta_N(k, \ell) &\geq \left(\frac{1}{2\pi i}\right)^2 \oint_{\partial\mathcal{R}} \oint_{\partial\mathcal{R}} \frac{\sigma^2 z_1 z_2 (1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2)) \tilde{v}_N(z_1, z_2) w'_N(z_1) w'_N(z_2)}{2(\lambda_{k,N} - w_N(z_1))(\lambda_{\ell,N} - w_N(z_1))(\lambda_{k,N} - w_N(z_2))(\lambda_{\ell,N} - w_N(z_2))} dz_1 dz_2 \\ &= \frac{\sigma^4}{2N} \sum_{m=1}^K \frac{1}{\lambda_{m,N}^2}. \end{aligned} \quad (85)$$

For $k \leq K$ and $\ell \geq K + 1$, we have

$$\vartheta_N(k, \ell) \geq \left(\frac{1}{2\pi i}\right)^2 \oint_{\partial\mathcal{R}} \oint_{\partial\mathcal{R}} \frac{\sigma^2 \lambda_{k,N} (1 - u_N(z_1, z_2)) \Delta_N(z_1, z_2)^{-1} w'_N(z_1) w'_N(z_2)}{2w_N(z_1)w_N(z_2)(\lambda_{k,N} - w_N(z_1))(\lambda_{k,N} - w_N(z_2))} dz_1 dz_2.$$

Using (80) and performing a serie expansion of $(1 - u_N(z_1, z_2))^{-1}$, we obtain as well

$$\vartheta_N(k, \ell) \geq \left(\frac{1}{2\pi i}\right)^2 \oint_{\partial\mathcal{R}} \oint_{\partial\mathcal{R}} \frac{\sigma^2 \lambda_{k,N} w'_N(z_1) w'_N(z_2)}{2w_N(z_1)w_N(z_2)(\lambda_{k,N} - w_N(z_1))(\lambda_{k,N} - w_N(z_2))} dz_1 dz_2 = \frac{\sigma^2}{2\lambda_{k,N}}. \quad (86)$$

(51) will follow from (84), (85) and (86) and lemma 2 in section 2.4.

4.4 Asymptotics of $\vartheta_N(k, l)$

We now show (52), by assuming that K is independent of N . In this case, using the results of section 2.4, it is not difficult to show that

$$\max_{k, \ell} |\vartheta_N(k, \ell) - \check{\vartheta}_N(k, \ell)| \xrightarrow{N \rightarrow \infty} 0, \quad (87)$$

where

$$\check{\vartheta}_N(k, \ell) = \left(\frac{1}{2\pi i}\right)^2 \oint_{\partial\mathcal{R}} \oint_{\partial\mathcal{R}} \frac{\frac{\sigma^2}{2} \left(\sigma^2 + \frac{\sigma^2 c_N \lambda_{k,N} \lambda_{\ell,N}}{w(z_1)w(z_2)} + (\lambda_{k,N} + \lambda_{\ell,N}) \right) \Delta(z_1, z_2)^{-1} w'(z_1) w'(z_2)}{(\lambda_{k,N} - w(z_1))(\lambda_{\ell,N} - w(z_1))(\lambda_{k,N} - w(z_2))(\lambda_{\ell,N} - w(z_2))} dz_1 dz_2 \quad (88)$$

with $m(z)$, $w(z)$ and $\Delta(z_1, z_2)$ defined in (26), (28) and (34). From (81), it is clear that the following serie expansion also holds,

$$\Delta(z_1, z_2)^{-1} = \left(1 - \frac{\sigma^4 c_N}{w(z_1)w(z_2)}\right)^{-1} = \sum_{n \in \mathbb{N}} \left(\frac{\sigma^4 c_N}{w(z_1)w(z_2)}\right)^n,$$

uniformly on $\partial\mathcal{R} \times \partial\mathcal{R}$, and thus

$$\begin{aligned} \check{\vartheta}_N(k, \ell) &= \sum_{n \in \mathbb{N}} \left(\frac{1}{2\pi i}\right)^2 \oint_{\partial\mathcal{R}} \oint_{\partial\mathcal{R}} \frac{\frac{\sigma^2}{2} \left(\sigma^2 + \frac{\sigma^2 c_N \lambda_{k,N} \lambda_{\ell,N}}{w(z_1)w(z_2)} + (\lambda_{k,N} + \lambda_{\ell,N}) \right) (\sigma^4 c_N)^n w(z_1)^{-n} w(z_2)^{-n} w'(z_1) w'(z_2)}{(\lambda_{k,N} - w(z_1))(\lambda_{\ell,N} - w(z_1))(\lambda_{k,N} - w(z_2))(\lambda_{\ell,N} - w(z_2))} dz_1 dz_2. \end{aligned}$$

By expressing the previous expression as square of single integrals, we obtain

$$\begin{aligned} \check{\vartheta}_N(k, \ell) &= \frac{\sigma^2}{2} (\sigma^2 + (\lambda_{k,N} + \lambda_{\ell,N})) \sum_{n \in \mathbb{N}} (\sigma^4 c_N)^n \left[\frac{1}{2\pi i} \oint_{w(\partial\mathcal{R})} \frac{1}{w^n (\lambda_{k,N} - w) (\lambda_{\ell,N} - w)} dw \right]^2 \\ &\quad + \frac{\sigma^4 c_N}{2} \lambda_{k,N} \lambda_{\ell,N} \sum_{n \in \mathbb{N}} (\sigma^4 c_N)^n \left[\frac{1}{2\pi i} \oint_{w(\partial\mathcal{R})} \frac{1}{w^{n+1} (\lambda_{k,N} - w) (\lambda_{\ell,N} - w)} dw \right]^2. \end{aligned}$$

Using classical residue computation, we eventually end up with (52),

$$\check{\vartheta}_N(k, \ell) = \frac{\sigma^4 c_N (\lambda_{k,N} \lambda_{\ell,N} + (\lambda_{k,N} + \lambda_{\ell,N}) \sigma^2 + \sigma^4) (\lambda_{k,N} \lambda_{\ell,N} + \sigma^4 c_N)}{2 (\lambda_{k,N}^2 - \sigma^4 c_N) (\lambda_{\ell,N}^2 - \sigma^4 c_N) (\lambda_{k,N} \lambda_{\ell,N} - \sigma^4 c_N)} (1 - \mathbb{1}_{[K+1, M]}(k) \mathbb{1}_{[K+1, M]}(\ell)). \quad (89)$$

We can easily show that

$$\liminf_{N \rightarrow \infty} \min_{(k, \ell) \notin \{K+1, \dots, M\}^2} \check{\vartheta}_N(k, \ell) > 0.$$

and we also have the boundedness

$$\limsup_{N \rightarrow \infty} \max_{k, \ell} \check{\vartheta}_N(k, \ell) < \infty,$$

which is ensured in the general case (K not necessarily fixed) by (79), but which also comes from lemma 3 in section 2.4.

4.5 Solution to the differential equation

Recall that the differential equation (76)

$$\psi'_N(u) = \left(i \sqrt{N} \operatorname{Re} (\xi_N (\mathbf{d}_{1,N}^* \mathbf{d}_{2,N} - \eta_N)) - u \boldsymbol{\xi}_N^T \boldsymbol{\Gamma}_N \boldsymbol{\xi}_N \right) \psi_N(u) + \frac{\epsilon_N(u)}{\sqrt{N}}, \quad (90)$$

where (ξ_N) is any deterministic sequence such that $\limsup_N |\xi_N| < \infty$ and $\boldsymbol{\xi}_N = [\operatorname{Re}(\xi_N), \operatorname{Im}(\xi_N)]^T$ and $u \mapsto \epsilon_N(u)$ a generic continuously differentiable function such that

$$\limsup_{N \rightarrow \infty} \{ |\epsilon_N(u)|, |\epsilon'_N(u)| \} \leq P(u),$$

with P is a polynomial independent of N with positive coefficient. By differentiating (90) with respect to u and using (64), one can check that

$$\mathbb{E} \left| \sqrt{N} \operatorname{Re} (\xi_N (\hat{\gamma}_N - (\mathbf{d}_{1,N}^* \mathbf{d}_{2,N} - \eta_N))) \right|^2 = \boldsymbol{\xi}_N^T \boldsymbol{\Gamma}_N \boldsymbol{\xi}_N + \mathcal{O} \left(\frac{1}{\sqrt{N}} \right),$$

where $\hat{\gamma}_N$ is defined in (63). This implies

$$\operatorname{Re} (\xi_N (\hat{\gamma}_N - (\mathbf{d}_{1,N}^* \mathbf{d}_{2,N} - \eta_N))) = \begin{cases} \mathcal{O}_{\mathbb{P}} \left(\frac{1}{\sqrt{N}} \right) & \text{if } \liminf_N \boldsymbol{\xi}_N^T \boldsymbol{\Gamma}_N \boldsymbol{\xi}_N > 0 \\ o_{\mathbb{P}} \left(\frac{1}{\sqrt{N}} \right) & \text{otherwise} \end{cases}.$$

which shows (53). By assuming that $\liminf_N \boldsymbol{\xi}_N^T \boldsymbol{\Gamma}_N \boldsymbol{\xi}_N > 0$, we can replace ξ_N by $\frac{\xi_N}{\sqrt{\boldsymbol{\xi}_N^T \boldsymbol{\Gamma}_N \boldsymbol{\xi}_N}}$ in (76) without modifying the boundedness property of $\epsilon_N(u)$, and thus

$$\psi'_N(u) = \left(i \frac{\sqrt{N} \operatorname{Re} (\xi_N (\mathbf{d}_{1,N}^* \mathbf{d}_{2,N} - \eta_N))}{\sqrt{\boldsymbol{\xi}_N^T \boldsymbol{\Gamma}_N \boldsymbol{\xi}_N}} - u \right) \psi_N(u) + \frac{\epsilon_N(u)}{\sqrt{N}}, \quad (91)$$

(91) being a classical nonhomogeneous linear differential equation of the first order, we easily obtain that

$$\mathbb{E} \left[\exp \left(i u \frac{\sqrt{N} \operatorname{Re} (\xi_N (\hat{\gamma}_N - (\mathbf{d}_{1,N}^* \mathbf{d}_{2,N} - \eta_N)))}{\sqrt{\boldsymbol{\xi}_N^T \boldsymbol{\Gamma}_N \boldsymbol{\xi}_N}} \right) \right] = \exp(-u^2/2) + o(1),$$

which proves (54).

5 Appendix

5.1 Proof of formula (66) and proposition 1

In this section, we prove formula (66) and proposition 1 respectively. For that purpose, we use two tools, an integration by part formula and a Poincaré inequality for Gaussian variables, which are well-known in the field of random matrix theory since the work of Pastur [13]. We first note that every function $f : \mathbb{C} \mapsto \mathbb{C}$ can be written as $f(z) = \tilde{f}(\operatorname{Re}(z), \operatorname{Im}(z))$. If $\tilde{f} \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{C})$, we define the usual differential operators

$$\frac{\partial f(z)}{\partial z} = \frac{1}{2} \left(\frac{\partial \tilde{f}(x, y)}{\partial x} - i \frac{\partial \tilde{f}(x, y)}{\partial y} \right) \quad \text{and} \quad \frac{\partial f(z)}{\partial \bar{z}} = \frac{1}{2} \left(\frac{\partial \tilde{f}(x, y)}{\partial x} + i \frac{\partial \tilde{f}(x, y)}{\partial y} \right).$$

In this context, we say that f is continuously differentiable if \tilde{f} is continuously differentiable. The following lemma gives the integration by part formula and the Poincaré inequality.

Lemma 4. *Let $z_1 = x_1 + iy_1, \dots, z_n = x_n + iy_n$ be n i.i.d. $\mathcal{N}_{\mathbb{C}}(0, \rho^2)$ variables and let f a continuously differentiable function defined on \mathbb{C}^n with polynomially bounded partial derivatives. Then, if $\mathbf{z} = (z_1, \dots, z_n)^T$, it holds that*

$$\mathbb{E}[z_k f(\mathbf{z})] = \rho^2 \mathbb{E} \left[\frac{\partial f(\mathbf{z})}{\partial \bar{z}_k} \right] \quad \text{and} \quad \mathbb{E}[\bar{z}_k f(\mathbf{z})] = \rho^2 \mathbb{E} \left[\frac{\partial f(\mathbf{z})}{\partial z_k} \right].$$

Moreover,

$$\mathbb{V}[f(\mathbf{z})] \leq \rho^2 \sum_{k=1}^n \left(\mathbb{E} \left| \frac{\partial f(\mathbf{z})}{\partial z_k} \right|^2 + \mathbb{E} \left| \frac{\partial f(\mathbf{z})}{\partial \bar{z}_k} \right|^2 \right). \quad (92)$$

Hereafter and in all the remainder of this appendix, $\epsilon_N(u, z_1, z_2)$ will denote a generic continuous function on $\mathbb{R} \times \partial\mathcal{R} \times \partial\mathcal{R}$ such that $u \mapsto \epsilon_N(u, z_1, z_2)$ is continuously differentiable and

$$\limsup_{N \rightarrow \infty} \sup_{(z_1, z_2) \in \partial\mathcal{R} \times \partial\mathcal{R}} \left\{ |\epsilon_N(u, z_1, z_2)|, \left| \frac{\partial \epsilon_N(u, z_1, z_2)}{\partial u} \right| \right\} < P(u),$$

with $P(u)$ a polynomial with positive coefficients. $\epsilon_N(u, z_1, z_2)$ may take different values from one line to another. We will also keep the notation $\epsilon_N(z_1, z_2)$ and $\epsilon_N(u, z_1)$ if $\epsilon_N(u, z_1, z_2)$ does not depend on u or z_2 .

We recall the quantity $\hat{\gamma}_N$, which is the regularized estimator defined in (63) by

$$\hat{\gamma}_N = \frac{1}{2\pi i} \oint_{\partial\mathcal{R}} \mathbf{d}_{1,N}^* \mathbf{Q}_N(z) \mathbf{d}_{2,N} \chi_N \frac{w'_N(z)}{1 + \sigma^2 c_N m_N(z)} dz.$$

Using lemma 4, it is not difficult to obtain, as in [7, Lem. 5.7], the following useful properties.

Corollary 3. *Let $(h_N)_{N \geq 1}$ be a sequence of continuously differentiable functions defined on $\mathbb{C}^{M(M+N)}$ with polynomially bounded partial derivatives satisfying the condition*

$$\limsup_{N \rightarrow \infty} \sup_{z \in \partial\mathcal{R}} |h_N(\operatorname{Vec}(\mathbf{Q}_N(z)), \operatorname{Vec}(\mathbf{\Sigma}_N)) \chi_N| < \infty. \quad (93)$$

Then, for all $k \in \mathbb{N}^*$, we have

$$\begin{aligned} & \mathbb{E} \left[h_N(\operatorname{Vec}(\mathbf{Q}_N(z)), \operatorname{Vec}(\mathbf{\Sigma}_N)) \chi_N e^{iu\sqrt{N}\operatorname{Re}(\hat{\gamma}_N)} \right] \\ &= \mathbb{E} \left[h_N(\operatorname{Vec}(\mathbf{Q}_N(z)), \operatorname{Vec}(\mathbf{\Sigma}_N)) \chi_N^k e^{iu\sqrt{N}\operatorname{Re}(\hat{\gamma}_N)} \right] + \frac{\epsilon_N(u, z)}{N^p}. \end{aligned} \quad (94)$$

for all $p \in \mathbb{N}$. Moreover,

$$\begin{aligned} & \mathbb{E} \left[W_{i,j} h_N(\operatorname{Vec}(\mathbf{Q}_N(z)), \operatorname{Vec}(\mathbf{\Sigma}_N)) \chi_N e^{iu\sqrt{N}\operatorname{Re}(\hat{\gamma}_N)} \right] = \\ & \frac{\sigma^2}{N} \mathbb{E} \left[\frac{\partial}{\partial \bar{W}_{i,j}} \left\{ h_N(\operatorname{Vec}(\mathbf{Q}_N(z)), \operatorname{Vec}(\mathbf{\Sigma}_N)) e^{iu\sqrt{N}\operatorname{Re}(\hat{\gamma}_N)} \right\} \chi_N \right] + \frac{\epsilon_N(u, z)}{N^p}, \end{aligned} \quad (95)$$

and

$$\begin{aligned} \mathbb{E} \left[\overline{W}_{i,j} h_N (\text{Vec} (\mathbf{Q}_N(z)), \text{Vec}(\mathbf{\Sigma}_N)) \chi_N e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] = \\ \frac{\sigma^2}{N} \mathbb{E} \left[\frac{\partial}{\partial W_{i,j}} \left\{ h_N (\text{Vec} (\mathbf{Q}_N(z)), \text{Vec}(\mathbf{\Sigma}_N)) e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right\} \chi_N \right] + \frac{\epsilon_N(u, z)}{N^p}. \end{aligned} \quad (96)$$

Finally, we have

$$\mathbb{E} \left[h_N (\text{Vec} (\mathbf{Q}_N(z)), \text{Vec}(\mathbf{\Sigma}_N)) D_{i,j} e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] = \frac{\epsilon_N(u, z)}{N^p}$$

and

$$\mathbb{E} \left[h_N (\text{Vec} (\mathbf{Q}_N(z)), \text{Vec}(\mathbf{\Sigma}_N)) D_{i,j}^* e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] = \frac{\epsilon_N(u, z)}{N^p},$$

for all $p \in \mathbb{N}$, with $D_{i,j} = \frac{\partial}{\partial W_{i,j}^*} \{\chi_N\}$.

We now introduce the matrix $\mathbf{R}_N(z)$ given by

$$\mathbf{R}_N(z) = \left(\frac{\mathbf{B}_N \mathbf{B}_N^*}{1 + \sigma^2 c_N \mathbb{E} [\hat{m}_N(z_1) \chi_N]} - z (1 + \sigma^2 c_N \mathbb{E} [\hat{m}_N(z_1) \chi_N]) + \sigma^2 (1 - c_N) \right)^{-1}.$$

Matrix $\mathbf{R}_N(z)$ is similar to $\mathbf{T}_N(z)$ where we just replaced $m_N(z)$ by $\mathbb{E}[\hat{m}_N(z) \chi_N]$. Since $\mathbb{E}[\hat{m}_N(z) \chi_N] - m_N(z) \rightarrow_N 0$, it is of course expected that $\mathbf{R}_N(z)$ will be close to $\mathbf{T}_N(z)$ asymptotically. This result is given by the following lemma.

Corollary 4 ([7, Lem. 3.10, 5.5 & 5.6]). *Let $\mathbf{M}_N(z)$ a sequence of deterministic matrices of size $M \times M$ such that*

$$\limsup_{N \rightarrow \infty} \sup_{z \in \partial \mathcal{R}} \|\mathbf{M}_N(z)\| < \infty.$$

Then we have

$$\limsup_{N \rightarrow \infty} \sup_{z \in \mathcal{R}} |1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z) \chi_N]|^{-1} < \infty \quad \text{and} \quad \limsup_{N \rightarrow \infty} \sup_{z \in \mathcal{R}} \|\mathbf{R}_N(z)\| < \infty.$$

Moreover, $\mathbb{E} [\hat{m}_N(z) \chi_N] - m_N(z) = \frac{\epsilon_N(z)}{N^2}$ and

$$\mathbf{d}_{1,N}^* (\mathbb{E}[\mathbf{Q}_N(z) \chi_N] - \mathbf{R}_N(z)) \mathbf{M}_N(z) \mathbf{d}_{2,N} = \frac{\epsilon_N(z)}{N^{3/2}} \quad \text{and} \quad \mathbf{d}_{1,N}^* (\mathbf{R}_N(z) - \mathbf{T}_N(z)) \mathbf{M}_N(z) \mathbf{d}_{2,N} = \frac{\epsilon_N(z)}{N^{3/2}}.$$

We now give a result on the variance of certain expressions involving the resolvent, whose proof is a standard application of the Poincaré inequality (see e.g. [7, Lem. 5.8], [16, Lem. 10]), and is therefore omitted.

Corollary 5. *Let $\mathbf{M}_N(z_1, z_2)$ a sequence of deterministic matrices of size $M \times M$ such that*

$$\limsup_{N \rightarrow \infty} \sup_{(z_1, z_2) \in \partial \mathcal{R}^2} \|\mathbf{M}_N(z_1, z_2)\| < \infty.$$

Let P a 5-variate polynomial function independent of N such that the matrix

$$\mathbf{\Xi}_N(z_1, z_2) = P(\mathbf{Q}_N(z_1), \mathbf{Q}_N(z_2), \mathbf{\Sigma}_N, \mathbf{\Sigma}_N^*, \mathbf{M}_N(z_1, z_2)) \chi_N.$$

is properly defined. Then, it holds that

$$\limsup_{N \rightarrow \infty} \sup_{(z_1, z_2) \in \partial \mathcal{R}^2} \mathbb{V} [\text{tr} \mathbf{\Xi}_N(z_1, z_2)] < \infty \quad \text{and} \quad \limsup_{N \rightarrow \infty} \sup_{(z_1, z_2) \in \partial \mathcal{R}^2} \mathbb{V} \left[\sqrt{N} \mathbf{d}_{1,N}^* \mathbf{\Xi}_N(z_1, z_2) \mathbf{d}_{2,N} \right] < \infty.$$

Moreover, it also holds that $\mathbb{V} \left[\mathbf{d}_{1,N}^* \mathbf{\Xi}_N(z_1, z_2) \mathbf{d}_{2,N} e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right]$ is a term behaving as $\epsilon_N(u, z_1, z_2)$.

5.1.1 Proof of formula (66)

By expressing the derivative of $e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)}$ w.r.t. $W_{i,j}$, we obtain

$$\begin{aligned} \frac{\partial}{\partial W_{i,j}} \left\{ e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right\} = \\ \frac{iu}{2} \frac{1}{2\pi i} \oint_{\partial\mathcal{R}} (-\mathbf{d}_{1,N}^* \mathbf{Q}_N(z) \mathbf{e}_i \mathbf{e}_j^* \Sigma_N^* \mathbf{Q}_N(z) \mathbf{d}_{2,N} \chi_N + \mathbf{d}_{1,N}^* \mathbf{Q}_N(z) \mathbf{d}_{2,N} D_{i,j}^*) \frac{w'_N(z)}{1 + \sigma^2 c_N m_N(z)} e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} dz \\ + \frac{iu}{2} \frac{1}{2\pi i} \oint_{\partial\mathcal{R}} (-\mathbf{d}_{2,N}^* \mathbf{Q}_N(z) \mathbf{e}_i \mathbf{e}_j^* \Sigma_N^* \mathbf{Q}_N(z) \mathbf{d}_{1,N} \chi_N + \mathbf{d}_{2,N}^* \mathbf{Q}_N(z) \mathbf{d}_{1,N} D_{i,j}^*) \frac{w'_N(z)}{1 + \sigma^2 c_N m_N(z)} e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} dz, \end{aligned} \quad (97)$$

where $D_{i,j}$ is defined in corollary 3. The derivative with respect to $\overline{W}_{i,j}$ is computed in the same way. To develop $\mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z) \mathbf{d}_{2,N} \chi_N e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right]$, we start with the classical resolvent identity $\mathbf{Q}_N(z) \Sigma_N \Sigma_N^* = \mathbf{I} + z \mathbf{Q}_N(z)$. By applying corollaries 3 and 4 several times, long but straightforward computations lead to

$$\begin{aligned} \mathbb{E} \left[[\mathbf{Q}_N(z_1) \Sigma_N \Sigma_N^*]_{i,j} \chi_N e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] = \\ \alpha_N^{(i,j)}(u, z_1) + \frac{iu\sigma^2}{2\sqrt{N}} \frac{1}{2\pi i} \oint_{\partial\mathcal{R}} \frac{\left(\beta_N^{(i,j)}(u, z_1, z_2) + \tilde{\beta}_N^{(i,j)}(u, z_1, z_2) \right) w'_N(z_2)}{(1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_1) \chi_N]) (1 + \sigma^2 c_N m_N(z_2))} dz_2 \\ - [\Delta_N(u, z_1)]_{i,j} + \frac{\epsilon_N(u, z_1)}{N^p}, \end{aligned} \quad (98)$$

for all $p \in \mathbb{N}$, where $\alpha_N^{(i,j)}(u, z_1)$ is given by

$$\begin{aligned} \alpha_N^{(i,j)}(u, z_1) = \\ \frac{\mathbb{E} \left[[\mathbf{Q}_N(z_1) \mathbf{B}_N \mathbf{B}_N^*]_{i,j} \chi_N e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right]}{1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_1) \chi_N]} + \frac{\sigma^2 \mathbb{E} \left[[\mathbf{Q}_N(z)]_{i,j} \chi_N e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right]}{1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_1) \chi_N]} \\ - \frac{\mathbb{E} \left[[\mathbf{Q}_N(z_1)]_{i,j} e^{iu\text{Re}(\hat{\gamma}_N)} \right]}{(1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_1) \chi_N])^2} \frac{\sigma^2}{N} \text{tr} \mathbf{B}_N^* \mathbb{E}[\mathbf{Q}_N(z_1) \chi_N] \mathbf{B}_N, \end{aligned}$$

$\beta_N^{(i,j)}(u, z_1, z_2)$ and $\tilde{\beta}_N^{(i,j)}(u, z_1, z_2)$ respectively by

$$\begin{aligned} \beta_N^{(i,j)}(u, z_1, z_2) = \\ \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_2) \mathbf{e}_j \mathbf{e}_i^* \mathbf{Q}_N(z_1) \mathbf{B}_N \Sigma_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N^2 e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] \\ + \mathbb{E} \left[\mathbf{d}_{2,N}^* \mathbf{Q}_N(z_2) \mathbf{e}_j \mathbf{e}_i^* \mathbf{Q}_N(z_1) \mathbf{B}_N \Sigma_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{1,N} \chi_N^2 e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] \end{aligned}$$

and

$$\begin{aligned} \tilde{\beta}_N^{(i,j)}(u, z_1, z_2) = \\ \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_2) \Sigma_N \Sigma_N^* \mathbf{e}_j \mathbf{e}_i^* \mathbf{Q}_N(z_1) \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N^2 e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] \\ + \mathbb{E} \left[\mathbf{d}_{2,N}^* \mathbf{Q}_N(z_2) \Sigma_N \Sigma_N^* \mathbf{e}_j \mathbf{e}_i^* \mathbf{Q}_N(z_1) \mathbf{Q}_N(z_2) \mathbf{d}_{1,N} \chi_N^2 e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] \end{aligned}$$

and where finally the matrix $\Delta_N(u, z_1)$ is given by

$$\begin{aligned} \Delta_N(u, z_1) = & - \frac{\mathbb{E} \left[\mathbf{Q}_N(z_1) \chi_N e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right]}{(1 + \sigma^2 c_N \mathbb{E} [\hat{m}_N(z_1) \chi_N])^2} \mathbb{E} \left[\left(\frac{\sigma^2}{N} \text{tr} \mathbf{Q}_N(z_1) \chi_N - \mathbb{E} \left[\frac{\sigma^2}{N} \text{tr} \mathbf{Q}_N(z_1) \chi_N \right] \right) \frac{\sigma^2}{N} \text{tr} \Sigma_N^* \mathbf{Q}_N(z_1) \mathbf{B}_N \chi_N \right] \\ & + \frac{1}{1 + \sigma^2 c_N \mathbb{E} [\hat{m}_N(z_1) \chi_N]} \mathbb{E} \left[\left(\frac{\sigma^2}{N} \text{tr} \Sigma_N^* \mathbf{Q}_N(z_1) \mathbf{B}_N \chi_N - \mathbb{E} \left[\frac{\sigma^2}{N} \text{tr} \Sigma_N^* \mathbf{Q}_N(z_1) \mathbf{B}_N \chi_N \right] \right) \mathbf{Q}_N(z) \chi_N e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] \\ & + \frac{1}{1 + \sigma^2 c_N \mathbb{E} [\hat{m}_N(z_1) \chi_N]} \mathbb{E} \left[\left(\frac{\sigma^2}{N} \text{tr} \mathbf{Q}_N(z_1) \chi_N - \mathbb{E} \left[\frac{\sigma^2}{N} \text{tr} \mathbf{Q}_N(z_1) \chi_N \right] \right) \mathbf{Q}_N(z_1) \Sigma_N \Sigma_N^* \chi_N e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right]. \end{aligned} \quad (99)$$

From (98) and the resolvent identity, we obtain

$$\begin{aligned} & \left[\mathbb{E} \left[\mathbf{Q}_N(z_1) \chi_N e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] \left(\frac{\mathbf{B}_N \mathbf{B}_N^*}{1 + \sigma^2 c_N \mathbb{E} [\hat{m}_N(z_1) \chi_N]} - z (1 + \sigma^2 c_N \tilde{\tau}_N(z_1)) \mathbf{I} \right) \right]_{i,j} = \\ & \mathbb{E} \left[[\mathbf{I}]_{i,j} \chi_N e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] - \frac{iu\sigma^2}{2\sqrt{N}} \frac{1}{2\pi i} \oint_{\partial \mathcal{R}} \frac{\left(\beta_N^{(i,j)}(u, z_1, z_2) + \tilde{\beta}_N^{(i,j)}(u, z_1, z_2) \right) w'_N(z_2)}{(1 + \sigma^2 c_N \mathbb{E} [\hat{m}_N(z_1) \chi_N]) (1 + \sigma^2 c_N m_N(z_2))} dz_2 \\ & + [\Delta_N(u, z_1)]_{i,j} + \frac{\epsilon_N(u, z_1)}{N^p}, \end{aligned}$$

where

$$\tilde{\tau}_N(z_1) = - \frac{\sigma}{z (1 + \sigma^2 c_N \mathbb{E} [\hat{m}_N(z_1) \chi_N])} \left(1 - \frac{1}{N} \text{tr} \left(\frac{\mathbf{B}_N^* \mathbb{E} [\mathbf{Q}_N(z_1)] \mathbf{B}_N}{1 + \sigma^2 c_N \mathbb{E} [\hat{m}_N(z_1) \chi_N]} \right) \right)$$

Straightforward algebra gives

$$\tilde{\tau}_N(z_1) = \mathbb{E} \left[\left(\hat{m}_N(z_1) - \frac{(1 - 1/c_N)}{z_1} \right) \chi_N \right] + \frac{1}{z_1} \frac{1}{M} \text{tr} \Delta_N(0, z_1) + \frac{\epsilon_N(z_1)}{N^p},$$

and finally we get

$$\begin{aligned} & \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{d}_{2,N} \chi_N e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] = \\ & \mathbf{d}_{1,N}^* \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \mathbb{E} \left[\chi_N e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] - \frac{iu\sigma^2}{2\sqrt{N}} \frac{1}{2\pi i} \oint_{\partial \mathcal{R}} \frac{\left(\beta_N(u, z_1, z_2) + \tilde{\beta}_N(u, z_1, z_2) \right) w'_N(z_2)}{(1 + \sigma^2 c_N \mathbb{E} [\hat{m}_N(z_1) \chi_N]) (1 + \sigma^2 c_N m_N(z_2))} dz_2 \\ & + \mathbf{d}_{1,N}^* \Delta_N(u, z_1) \mathbf{R}_N(z_2) \mathbf{d}_{2,N} + \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{d}_{2,N} e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] \frac{\sigma^2}{N} \text{tr} \Delta_N(0, z_1) + \frac{\epsilon_N(u, z_1)}{N^p}. \end{aligned} \quad (100)$$

where $\beta_N(u, z_1, z_2)$ is defined by

$$\begin{aligned} \beta_N(u, z_1, z_2) = & \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_2) \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{B}_N \Sigma_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N^2 e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] \\ & + \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_2) \Sigma_N \Sigma_N^* \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N^2 e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right], \end{aligned}$$

and $\tilde{\beta}_N(u, z_1, z_2)$ by

$$\begin{aligned} \tilde{\beta}_N(u, z_1, z_2) = & \mathbb{E} \left[\mathbf{d}_{2,N}^* \mathbf{Q}_N(z_2) \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{B}_N \Sigma_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{1,N} \chi_N^2 e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right] \\ & + \mathbb{E} \left[\mathbf{d}_{2,N}^* \mathbf{Q}_N(z_2) \Sigma_N \Sigma_N^* \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{Q}_N(z_2) \mathbf{d}_{1,N} \chi_N^2 e^{iu\sqrt{N}\text{Re}(\hat{\gamma}_N)} \right]. \end{aligned}$$

Using corollaries 4 and 5 (in conjunction with Cauchy-Schwarz inequality), and the fact that $\mathbb{E}[\chi_N e^{iu \operatorname{Re}(\hat{\gamma}_N)}] = \psi_N(u) + \mathcal{O}(N^{-p})$ (by dominated convergence theorem, see section 4.1), it is straightforward to show that

$$\begin{aligned} \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{d}_{2,N} \chi_N e^{iu \sqrt{N} \operatorname{Re}(\hat{\gamma}_N)} \right] = \\ \mathbf{d}_{1,N}^* \mathbf{T}_N(z_1) \mathbf{d}_{2,N} \psi_N(u) - \frac{iu\sigma^2}{2\sqrt{N}} \frac{1}{2\pi i} \oint_{\partial R} \frac{\left(\beta_N(u, z_1, z_2) + \tilde{\beta}_N(u, z_1, z_2) \right) w'_N(z_2)}{(1 + \sigma^2 c_N m_N(z_1)) (1 + \sigma^2 c_N m_N(z_2))} dz_2 + \frac{\epsilon_N(u, z_1)}{N}. \end{aligned} \quad (101)$$

The next step consists in decorrelating the different terms inside the expressions of β_N and $\tilde{\beta}_N$. We have

$$\begin{aligned} \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_2) \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{B}_N \Sigma_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N^2 e^{iu \sqrt{N} \operatorname{Re}(\hat{\gamma}_N)} \right] = \\ \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_2) \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \chi_N \right] \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{B}_N \Sigma_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N \right] \psi_N(u) \\ + \mathbb{E} \left[\mathbf{d}_{1,N}^* (\mathbf{Q}_N(z_2) - \mathbb{E}[(\mathbf{Q}_N(z_2))]) \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{B}_N \Sigma_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N^2 e^{iu \sqrt{N} \operatorname{Re}(\hat{\gamma}_N)} \right] \\ + \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_2) \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \chi_N \right] \mathbb{E} \left[\mathbf{d}_{1,N}^* (\mathbf{Q}_N(z_1) \mathbf{B}_N \Sigma_N^* \mathbf{Q}_N(z_2) - \mathbb{E}[\mathbf{Q}_N(z_1) \mathbf{B}_N \Sigma_N^* \mathbf{Q}_N(z_2)]) \mathbf{d}_{2,N} \chi_N^2 e^{iu \sqrt{N} \operatorname{Re}(\hat{\gamma}_N)} \right], \end{aligned}$$

and using again corollaries 4 and 5, we end up with

$$\begin{aligned} \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_2) \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{B}_N \Sigma_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N^2 e^{iu \sqrt{N} \operatorname{Re}(\hat{\gamma}_N)} \right] = \\ \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_2) \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \chi_N \right] \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{B}_N \Sigma_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N \right] \psi_N(u) + \frac{\epsilon_N(u, z_1, z_2)}{\sqrt{N}}. \end{aligned}$$

In the same way,

$$\begin{aligned} \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_2) \Sigma_N \Sigma_N^* \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N^2 e^{iu \sqrt{N} \operatorname{Re}(\hat{\gamma}_N)} \right] = \\ \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_2) \Sigma_N \Sigma_N^* \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \chi_N \right] \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N \right] \psi_N(u) + \frac{\epsilon_N(u, z_1, z_2)}{\sqrt{N}}. \end{aligned}$$

A standard application of corollaries 3 4 and 5 leads to

$$\begin{aligned} \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{B}_N \Sigma_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N \right] = \\ \frac{\mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{B}_N \mathbf{B}_N^* \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \chi_N \right]}{1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_2)]} - \frac{\mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_1) \mathbf{Q}_N(z_2) \mathbf{d}_{2,N} \right] \frac{\sigma^2}{N} \operatorname{tr} \mathbf{B}_N^* \mathbb{E}[\mathbf{Q}_N(z_1) \chi_N] \mathbf{B}_N}{(1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_1)]) (1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_2)])} \\ + \frac{\epsilon_N(u, z_1, z_2)}{N} \end{aligned}$$

and

$$\begin{aligned} \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_2) \Sigma_N \Sigma_N^* \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \chi_N \right] = \\ \frac{\mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_2) \mathbf{B}_N \mathbf{B}_N^* \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \right]}{1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_2)]} \\ - \left(\frac{\frac{\sigma^2}{N} \operatorname{tr} \mathbf{B}_N^* \mathbb{E}[\mathbf{Q}_N(z_2)] \mathbf{B}_N}{(1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_2)])^2} + \frac{\sigma^2}{1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_2)]} \right) \mathbb{E} \left[\mathbf{d}_{1,N}^* \mathbf{Q}_N(z_2) \mathbf{R}_N(z_1) \mathbf{d}_{2,N} \right] + \frac{\epsilon_N(u, z_1, z_2)}{N^{3/2}}. \quad (102) \end{aligned}$$

Inserting the previous estimates into the expressions of $\beta_N(u, z_1, z_2)$ and $\tilde{\beta}_N(u, z_1, z_2)$, and replacing $\mathbb{E}[\hat{m}_N(z)]$ by $m_N(z)$ as well as $\mathbb{E}[\mathbf{Q}_N(z)]$ and $\mathbf{R}_N(z)$ by $\mathbf{T}_N(z)$ thanks to corollary 4, we finally obtain (66).

5.1.2 Proof of proposition 1

Since the proof of proposition 1 uses the same technic as in the proof of formula (66) (see appendix 5.1.1), we will only provide the main lines of the computations.

Let $\mathbf{M}_N(z_1, z_2)$ be a $M \times M$ deterministic matrix s.t.

$$\limsup_{N \rightarrow \infty} \sup_{(z_1, z_2) \in \partial \mathcal{R} \times \partial \mathcal{R}} \|\mathbf{M}_N(z_1, z_2)\| < \infty.$$

We will also use the generic notation $\mathbf{E}_N(z_1, z_2)$ for $M \times M$ matrices such that

$$\limsup_{N \rightarrow \infty} \sup_{(z_1, z_2) \in \partial \mathcal{R} \times \partial \mathcal{R}} |\text{tr } \mathbf{E}_N(z_1, z_2)| < \infty,$$

i.e. such that $\text{tr } \mathbf{E}_N(z_1, z_2)$ behaves as $\epsilon_N(z_1, z_2)$. The value of $\mathbf{E}_N(z_1, z_2)$ may change from one line to another.

Starting from the matrix $\mathbb{E}[\mathbf{Q}_N(z_1)\mathbf{M}_N(z_1, z_2)\mathbf{Q}_N(z_2)\chi_N]$, a repeated use of corollaries 3 4 and 5 together with the decorrelation trick, in the same way as in appendix 5.1.1, leads to

$$\begin{aligned} \mathbb{E} \left[\frac{\mathbf{Q}_N(z_1)\mathbf{M}_N(z_1, z_2)\mathbf{Q}_N(z_2)}{z_2(1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_2)])} \left(z_2(1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_2)]) \mathbf{I} - \mathbf{B}_N \mathbf{B}_N^* - \sigma^2 \left(\mathbf{I} - \frac{\frac{\sigma^2}{N} \text{tr } \mathbf{B}_N^* \mathbb{E}[\mathbf{Q}_N(z_2)] \mathbf{B}_N}{1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_2)]} \right) \right) \chi_N \right] = \\ - \frac{1}{z_2} \mathbb{E}[\mathbf{Q}_N(z_1)\mathbf{M}_N(z_1)] \chi_N - \frac{\mathbb{E}[\mathbf{Q}_N(z_1)\Sigma_N \Sigma_N^* \chi_N]}{z_2(1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_2)])} \mathbb{E} \left[\frac{\sigma^2}{N} \text{tr } \mathbf{Q}_N(z_1)\mathbf{M}_N(z_1, z_2)\mathbf{Q}_N(z_2)\chi_N \right] \\ - \frac{\mathbb{E}[\mathbf{Q}_N(z_1)\chi_N]}{z_2(1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_1)])} \frac{1}{(1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_2)])} \mathbb{E} \left[\frac{\sigma^2}{N} \text{tr } \mathbf{Q}_N(z_1)\mathbf{M}_N(z_1, z_2)\mathbf{Q}_N(z_2)\mathbf{B}_N \mathbf{B}_N^* \chi_N \right] \\ + \frac{\mathbb{E}[\mathbf{Q}_N(z_1)\chi_N] \frac{\sigma^2}{N} \text{tr } \mathbf{B}_N^* \mathbb{E}[\mathbf{Q}_N(z_2)] \mathbf{B}_N}{z_2(1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_1)])} \frac{1}{(1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_2)])^2} \mathbb{E} \left[\frac{\sigma^2}{N} \text{tr } \mathbf{Q}_N(z_1)\mathbf{M}_N(z_1, z_2)\mathbf{Q}_N(z_2)\chi_N \right] \\ + \frac{\mathbf{E}_N(z_1, z_2)}{N^2}. \end{aligned}$$

By introducing the matrix $\mathbf{R}_N(z_2)$ and as in appendix 5.1.1, we obtain

$$\begin{aligned} \mathbb{E}[\mathbf{Q}_N(z_1)\mathbf{M}_N(z_1, z_2)\mathbf{Q}_N(z_2)] = \\ \mathbb{E}[\mathbf{Q}_N(z_1)\mathbf{M}_N(z_1)\mathbf{R}_N(z_2)\chi_N] \chi_N + \frac{\mathbb{E}[\mathbf{Q}_N(z_1)\Sigma_N \Sigma_N^* \mathbf{R}_N(z_2)\chi_N]}{1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_2)]} \mathbb{E} \left[\frac{\sigma^2}{N} \text{tr } \mathbf{Q}_N(z_1)\mathbf{M}_N(z_1, z_2)\mathbf{Q}_N(z_2)\chi_N \right] \\ + \frac{\mathbb{E}[\mathbf{Q}_N(z_1)\chi_N]}{(1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_1)])} \frac{1}{(1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_2)])} \mathbb{E} \left[\frac{\sigma^2}{N} \text{tr } \mathbf{Q}_N(z_1)\mathbf{M}_N(z_1, z_2)\mathbf{Q}_N(z_2)\mathbf{B}_N \mathbf{B}_N^* \chi_N \right] \\ - \frac{\mathbb{E}[\mathbf{Q}_N(z_1)\chi_N] \frac{\sigma^2}{N} \text{tr } \mathbf{B}_N^* \mathbb{E}[\mathbf{Q}_N(z_2)] \mathbf{B}_N}{(1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_1)])} \frac{1}{(1 + \sigma^2 c_N \mathbb{E}[\hat{m}_N(z_2)])^2} \mathbb{E} \left[\frac{\sigma^2}{N} \text{tr } \mathbf{Q}_N(z_1)\mathbf{M}_N(z_1, z_2)\mathbf{Q}_N(z_2)\chi_N \right] \\ + \frac{\mathbf{E}_N(z_1, z_2)}{N^2}, \end{aligned}$$

By taking the trace in the previous expression, and using corollaries 3 4 and 5, we end up with the following 2×2 linear system

$$\begin{aligned} \begin{bmatrix} \mathbb{E} \left[\frac{\sigma^2}{N} \text{tr } \mathbf{Q}_N(z_1)\mathbf{M}_N(z_1, z_2)\mathbf{Q}_N(z_2)\chi_N \right] \\ \mathbb{E} \left[\frac{\sigma^2}{N} \text{tr } \frac{\mathbf{Q}_N(z_1)\mathbf{M}_N(z_1, z_2)\mathbf{Q}_N(z_2)\mathbf{B}_N \mathbf{B}_N^*}{(1 + \sigma^2 c_N m_N(z_1))(1 + \sigma^2 c_N m_N(z_2))} \chi_N \right] \end{bmatrix} = \\ \begin{bmatrix} v_N(z_1, z_2)s_N(z_1, z_2) + u_N(z_1, z_2) & v_N(z_1, z_2) \\ u_N(z_1, z_2)s_N(z_1, z_2) + r_N(z_1, z_2) & u_N(z_1, z_2) \end{bmatrix} \begin{bmatrix} \mathbb{E} \left[\frac{\sigma^2}{N} \text{tr } \mathbf{Q}_N(z_1)\mathbf{M}_N(z_1, z_2)\mathbf{Q}_N(z_2)\chi_N \right] \\ \mathbb{E} \left[\frac{\sigma^2}{N} \text{tr } \frac{\mathbf{Q}_N(z_1)\mathbf{M}_N(z_1, z_2)\mathbf{Q}_N(z_2)\mathbf{B}_N \mathbf{B}_N^*}{(1 + \sigma^2 c_N m_N(z_1))(1 + \sigma^2 c_N m_N(z_2))} \chi_N \right] \end{bmatrix} \\ + \begin{bmatrix} \frac{\sigma^2}{N} \text{tr } \mathbf{T}_N(z_1)\mathbf{M}_N(z_1, z_2)\mathbf{T}_N(z_2) \\ \frac{\sigma^2}{N} \text{tr } \mathbf{T}_N(z_1)\mathbf{M}_N(z_1, z_2)\mathbf{T}_N(z_2)\mathbf{B}_N \mathbf{B}_N^* \end{bmatrix} + \frac{1}{N^2} \begin{bmatrix} \epsilon_N(z_1, z_2) \\ \epsilon_N(z_1, z_2) \end{bmatrix}, \quad (103) \end{aligned}$$

where $u_N(z_1, z_2), v_N(z_1, z_2), r_N(z_1, z_2)$ and $s_N(z_1, z_2)$ are respectively defined in (16), (17), (70) and (69). The determinant of the previous system is given by

$$\Delta_N(z_1, z_2) = (1 - u_N(z_1, z_2))^2 - v_N(z_1, z_2)(s_N(z_1, z_2) + r_N(z_1, z_2)).$$

By relating $\mathbf{T}_N(z)$ with $\tilde{\mathbf{T}}_N(z)$, we obtain the equality

$$\frac{\mathbf{B}_N^* \mathbf{T}_N(z) \mathbf{B}_N}{1 + \sigma^2 c_N m_N(z)} = \mathbf{I} + z (1 + \sigma^2 c_N m_N(z)) \tilde{\mathbf{T}}_N(z). \quad (104)$$

Inserting relation (104) in the expressions of $r_N(z_1, z_2)$ and $s_N(z_1, z_2)$, we obtain respectively

$$s_N(z_1, z_2) = \frac{-\sigma^2}{(1 + \sigma^2 c_N m_N(z_1))(1 + \sigma^2 c_N m_N(z_2))} - z_1 \frac{\sigma^2}{N} \text{tr} \frac{\mathbf{T}_N(z_1)}{1 + \sigma^2 c_N m_N(z_2)} - z_2 \frac{\sigma^2}{N} \text{tr} \frac{\mathbf{T}_N(z_2)}{1 + \sigma^2 c_N m_N(z_1)}$$

and

$$r_N(z_1, z_2) = z_1 z_2 \tilde{v}_N(z_1, z_2) - \frac{-\sigma^2}{(1 + \sigma^2 c_N m_N(z_1))(1 + \sigma^2 c_N m_N(z_2))} + z_1 \frac{\sigma^2}{N} \text{tr} \frac{\mathbf{T}_N(z_1)}{1 + \sigma^2 c_N m_N(z_2)} + z_2 \frac{\sigma^2}{N} \text{tr} \frac{\mathbf{T}_N(z_2)}{1 + \sigma^2 c_N m_N(z_1)},$$

where $\tilde{v}_N(z_1, z_2)$ is defined by (17). The determinant thus writes

$$\Delta_N(z_1, z_2) = (1 - u_N(z_1, z_2))^2 - z_1 z_2 v_N(z_1, z_2) \tilde{v}_N(z_1, z_2).$$

Using lemma 1, we can finally solve the system (103) to obtain

$$\mathbb{E} \left[\frac{\sigma^2}{N} \text{tr} \mathbf{Q}_N(z_1) \mathbf{M}_N(z_1, z_2) \mathbf{Q}_N(z_2) \chi_N \right] = \frac{(1 - u_N(z_1, z_2)) \frac{\sigma^2}{N} \text{tr} \mathbf{T}_N(z_1) \mathbf{M}_N(z_1, z_2) \mathbf{T}_N(z_2) + v_N(z_1, z_2) \frac{\sigma^2}{N} \text{tr} \frac{\mathbf{T}_N(z_1) \mathbf{M}_N(z_1, z_2) \mathbf{T}_N(z_2) \mathbf{B}_N \mathbf{B}_N^*}{(1 + \sigma^2 c_N m_N(z_1))(1 + \sigma^2 c_N m_N(z_2))}}{\Delta_N(z_1, z_2)} + \frac{\epsilon_N(z_1, z_2)}{N^2}$$

and

$$\begin{aligned} \mathbb{E} \left[\frac{\sigma^2}{N} \text{tr} \frac{\mathbf{Q}_N(z_1) \mathbf{M}_N(z_1, z_2) \mathbf{Q}_N(z_2) \mathbf{B}_N \mathbf{B}_N^*}{(1 + \sigma^2 c_N m_N(z_1))(1 + \sigma^2 c_N m_N(z_2))} \chi_N \right] = \\ \frac{(u_N(z_1, z_2) s_N(z_1, z_2) + r_N(z_1, z_2)) \frac{\sigma^2}{N} \text{tr} \mathbf{T}_N(z_1) \mathbf{M}_N(z_1, z_2) \mathbf{T}_N(z_2)}{\Delta_N(z_1, z_2)} \\ + \frac{(1 - v_N(z_1, z_2) s_N(z_1, z_2) - u_N(z_1, z_2)) \frac{\sigma^2}{N} \text{tr} \frac{\mathbf{T}_N(z_1) \mathbf{M}_N(z_1, z_2) \mathbf{T}_N(z_2) \mathbf{B}_N \mathbf{B}_N^*}{(1 + \sigma^2 c_N m_N(z_1))(1 + \sigma^2 c_N m_N(z_2))}}{\Delta_N(z_1, z_2)} + \frac{\epsilon_N(z_1, z_2)}{N^2}. \end{aligned}$$

The approximations (71) and (72) will unfold by choosing $\mathbf{M}_N(z_1, z_2) = \mathbf{d}_{1,N} \mathbf{d}_{2,N}^*$, which concludes the proof of proposition 1.

5.2 Proof of lemma 1

We recall here that $\Delta_N(z_1, z_2)$ is defined (18) by

$$\Delta_N(z_1, z_2) = (1 - u_N(z_1, z_2))^2 - z_1 z_2 v_N(z_1, z_2) \tilde{v}_N(z_1, z_2),$$

with $u_N(z_1, z_2)$, $v_N(z_1, z_2)$ and $\tilde{v}_N(z_1, z_2)$ given by (16) and (17). We also recall the following bounds from [6] :

$$\sup_{N \rightarrow \infty} \sup_{z \in \mathcal{K}} |u_N(z, z^*)| < 1 \quad \text{and} \quad \inf_{N \rightarrow \infty} \inf_{z \in \mathcal{K}} |\Delta_N(z, z^*)| > 0. \quad (105)$$

For $z_1, z_2 \in \partial \mathcal{K}$, Cauchy-Schwarz inequality and (105) gives

$$|u_N(z_1, z_2)| \leq |u_N(z_1, z_1^*)|^{1/2} |u_N(z_2, z_2^*)|^{1/2}, \quad (106)$$

and thus

$$\limsup_{N \rightarrow \infty} \sup_{z_1, z_2 \in \mathcal{K}} |u_N(z_1, z_2)| < 1,$$

which shows (20). In the same way, since

$$|\Delta_N(z_1, z_2)| \geq (1 - |u_N(z_1, z_2)|)^2 - |z_1||z_2| |v_N(z_1, z_2)| |\tilde{v}_N(z_1, z_2)|,$$

a straightforward application of Cauchy-Schwarz inequality yields

$$|\Delta_N(z_1, z_2)| \geq \left(1 - |u_N(z_1, z_1^*)|^{1/2} |u_N(z_2, z_2^*)|^{1/2}\right)^2 - |z_1||z_2| |v_N(z_1, z_1^*)|^{1/2} |v_N(z_2, z_2^*)|^{1/2} |\tilde{v}_N(z_1, z_1^*)|^{1/2} |\tilde{v}_N(z_2, z_2^*)|^{1/2}.$$

Using the inequality

$$(1 - \sqrt{x_1 x_2})^2 \geq \sqrt{(1 - x_1)^2 - s_1 t_1} \sqrt{(1 - x_2)^2 - s_2 t_2},$$

valid for $x_i \in [0, 1]$ and $(1 - x_i)^2 \geq s_i t_i$ ($i = 1, 2$), we finally obtain

$$|\Delta_N(z_1, z_2)| \geq |\Delta_N(z_1, z_1^*)|^{1/2} |\Delta_N(z_2, z_2^*)|^{1/2}, \quad (107)$$

which readily implies

$$\inf_{N \rightarrow \infty} \inf_{z_1, z_2 \in \mathcal{K}} |\Delta_N(z_1, z_2)| > 0.$$

Moreover, from (12), and the definition of $u_N(z_1, z_2)$, $v_N(z_1, z_2)$ and $\tilde{v}_N(z_1, z_2)$, we also see that

$$|\Delta_N(z_1, z_2)| \leq P \left(\frac{1}{d(z_1, \text{supp}(\mu_N) \cup \{0\})} \right) Q \left(\frac{1}{d(z_2, \text{supp}(\mu_N) \cup \{0\})} \right), \quad (108)$$

where P, Q are two polynomials independent of N, z_1, z_2 with positive coefficients, and we thus deduce

$$\sup_{N \rightarrow \infty} \sup_{z_1, z_2 \in \mathcal{K}} |\Delta_N(z_1, z_2)| < \infty,$$

which shows (21). We now prove (19). By straightforward computations, it is easily shown that $\Delta_N(z_1, z_2)$ is the determinant of the following 2×2 linear system

$$\begin{bmatrix} \sigma c_N(m_N(z_1) - m_N(z_2)) \\ \sigma(z_1 \tilde{m}_N(z_1) - z_2 \tilde{m}_N(z_2)) \end{bmatrix} = \begin{bmatrix} u_N(z_1, z_2) & v_N(z_1, z_2) \\ z_1 z_2 \tilde{v}_N(z_1, z_2) & u_N(z_1, z_2) \end{bmatrix} \begin{bmatrix} \sigma c_N(m_N(z_1) - m_N(z_2)) \\ \sigma(z_1 \tilde{m}_N(z_1) - z_2 \tilde{m}_N(z_2)) \end{bmatrix} + \frac{z_1 - z_2}{\sigma} \begin{bmatrix} v_N(z_1, z_2) \\ u_N(z_1, z_2) \end{bmatrix},$$

Since $\Delta_N(z_1, z_2) \neq 0$ for $z_1, z_2 \in \mathcal{K}$, solving the previous linear system gives

$$\sigma c_N(m_N(z_1) - m_N(z_2)) = \frac{z_1 - z_2}{\Delta_N(z_1, z_2)} \frac{v_N(z_1, z_2)}{\sigma},$$

and it is easy to show that

$$\sigma c_N(m_N(z_1) - m_N(z_2)) = (w_N(z_1) - w_N(z_2)) \frac{v_N(z_1, z_2)}{\sigma}.$$

Thus we obtain the relation

$$w_N(z_1) - w_N(z_2) = \frac{z_1 - z_2}{\Delta_N(z_1, z_2)}$$

which shows (19). Finally, to prove (22), we write, using Cauchy-Schwarz inequality,

$$\begin{aligned} \left| \Delta_N(z_1, z_2) - (1 - u_N(z_1, z_2))^2 \right| &= |z_1||z_2| |v_N(z_1, z_2) \tilde{v}_N(z_1, z_2)| \\ &\leq |z_1||z_2| v_N(z_1, z_1^*)^{1/2} \tilde{v}_N(z_2, z_2^*)^{1/2} v_N(z_1, z_1^*)^{1/2} \tilde{v}_N(z_2, z_2^*)^{1/2}, \end{aligned}$$

and since $(1 - u_N(z, z^*))^2 > |z|^2 v_N(z, z^*) \tilde{v}_N(z, z^*)$ for $z \in \mathcal{K}$, we have

$$\left| \Delta_N(z_1, z_2) - (1 - u_N(z_1, z_2))^2 \right| < (1 - u_N(z_1, z_1^*)) (1 - u_N(z_2, z_2^*)).$$

As $|1 - u_N(z_1, z_2)| \geq 1 - |u_N(z_1, z_1^*)|^{1/2} |u_N(z_2, z_2^*)|^{1/2}$ and thanks to the inequality $\sqrt{(1-a)(1-b)} \leq 1 - \sqrt{ab}$ valid for $a, b \in [0, 1]$, we finally obtain

$$\left| \Delta_N(z_1, z_2) - (1 - u_N(z_1, z_2))^2 \right| < |1 - u_N(z_1, z_2)|^2,$$

or equivalently

$$\left| \frac{\Delta_N(z_1, z_2)}{(1 - u_N(z_1, z_2))^2} - 1 \right| < 1.$$

5.3 Proof of lemma 3

Assume that the separation condition **A-1** and **A-2** hold, and let $t \in (t_1^+, t_2^-)$. Since μ_N converges to the Marchenko-Pastur distribution, $w_N(t) > 0 \rightarrow w(t) \geq 0$ and we deduce that $t > \sigma^2(1 + \sqrt{c})^2$. From (29) and the behaviour of ϕ , we have $w(t) > \sigma^2\sqrt{c}$, and finally assumption **A-2** implies that $\liminf_N \lambda_{K,N} \geq w(t)$, which proves (31).

Now, assume that (31) holds and let $\epsilon > 0$ such that

$$\liminf_{N \rightarrow \infty} \lambda_{K,N} > \sigma^2\sqrt{c} + \epsilon. \quad (109)$$

For any compact $\mathcal{K} \subset (-\infty, 0) \cup (0, \sigma^2\sqrt{c} + \epsilon)$, we have

$$\sup_{w \in \mathcal{K}} \left| f_N(w) + \frac{1}{w} \right| \xrightarrow{N \rightarrow \infty} 0, \quad \sup_{w \in \mathcal{K}} |\phi_N(w) - \phi(w)| \xrightarrow{N \rightarrow \infty} 0, \quad \text{and} \quad \sup_{w \in \mathcal{K}} |\phi'_N(w) - \phi'(w)| \xrightarrow{N \rightarrow \infty} 0 \quad (110)$$

Since $\phi(w)$ has a unique maximum $\sigma^2(1 - \sqrt{c})^2$ at point $w = -\sigma^2\sqrt{c}$ on the interval $(-\infty, 0)$, ϕ_N will also admit a positive maximum in this interval for all large N , and thus

$$x_{1,N}^- = \sigma^2(1 - \sqrt{c})^2 + o(1) \quad \text{and} \quad w_N(x_{1,N}^-) = -\sigma^2\sqrt{c} + o(1). \quad (111)$$

In the same way, ϕ has a unique positive minimum $\sigma^2(1 + \sqrt{c})^2$ at $\sigma^2\sqrt{c}$ on the interval $(0, \sigma^2\sqrt{c} + \epsilon)$, and thus ϕ_N will also admit a positive minimum on this interval, at the point $w_N(x_{1,N}^+)$, for N large enough, and

$$x_{1,N}^+ = \sigma^2(1 + \sqrt{c})^2 + o(1) \quad \text{and} \quad w_N(x_{1,N}^+) = \sigma^2\sqrt{c} + o(1). \quad (112)$$

Therefore, we can find t_1^- such that $\liminf_N x_{1,N}^- > t_1^- > 0$. Moreover, if \mathcal{K}' is a compact included in $(\sigma^2\sqrt{c}, \sigma^2\sqrt{c} + \epsilon)$, (110) also implies that for N large enough,

$$\inf_{w \in \mathcal{K}'} 1 - \sigma^2 c_N f_N(w) > 0 \quad \text{and} \quad \inf_{w \in \mathcal{K}'} \phi'_N(w) > 0. \quad (113)$$

which proves that $\phi_N(\mathcal{K}') \subset \mathbb{R} \setminus \text{supp}(\mu_N)$ from [16, Lemma 6]. This shows that it necessarily exists a local maximum $x_{2,N}^-$ of ϕ_N with preimage $w_N(x_{2,N}^-) > \sup \mathcal{K}'$. By fixing two points $t_2^- > t_1^+ > \sigma^2(1 + \sqrt{c})^2$ such that $w(t_1^+), w(t_2^-) \in \text{Int}(\mathcal{K}')$, we easily conclude that

$$\limsup_{N \rightarrow \infty} x_{1,N}^+ < t_1^+ \quad \text{and} \quad \liminf_{N \rightarrow \infty} x_{2,N}^- > t_2^-, \quad (114)$$

which proves **A-1**. By definition of t_2^- , $w(t_2^-) < \sigma^2\sqrt{c} + \epsilon$, which implies of course **A-2**.

References

- [1] T.W. Anderson. *An introduction to multivariate statistical analysis*, volume 2. Wiley New York, 1958.
- [2] F. Benaych-Georges and R.R. Nadakuditi. The singular values and vectors of low rank perturbations of large rectangular random matrices. *To appear in Journal of Multivariate Analysis*, 2011.

- [3] T. Citron and T. Kailath. An improved eigenvector beamformer. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'84.*, volume 9, pages 718–721. IEEE, 1984.
- [4] R.B. Dozier and J.W. Silverstein. Analysis of the limiting spectral distribution of large dimensional information-plus-noise type matrices. *Journal of Multivariate Analysis*, 98(6):1099–1122, 2007.
- [5] R.B. Dozier and J.W. Silverstein. On the empirical distribution of eigenvalues of large dimensional information-plus-noise-type matrices. *Journal of Multivariate Analysis*, 98(4):678–694, 2007.
- [6] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet. A subspace estimator for fixed rank perturbations of large random matrices. *Journal of Multivariate Analysis*, 114:427–447. arXiv:1106.1497.
- [7] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet. Large information plus noise random matrix models and consistent subspace estimation in large sensor networks. *Random Matrices: Theory and Applications*, 1(2), 2012.
- [8] H. Liu and G. Xu. A subspace method for signature waveform estimation in synchronous cdma systems. *Communications, IEEE Transactions on*, 44(10):1346–1354, 1996.
- [9] Philippe Loubaton and Pascal Vallet. Almost sure localization of the eigenvalues in a gaussian information plus noise model. application to the spiked models. *Electron. J. Probab.*, 16:1934–1959, 2011.
- [10] V.A. Marchenko and L.A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1:457, 1967.
- [11] X. Mestre. Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates. *Information Theory, IEEE Transactions on*, 54(11):5113–5129, 2008.
- [12] X. Mestre, P. Vallet, P. Loubaton, and W. Hachem. Asymptotic analysis of a consistent subspace estimator for observations of increasing dimension. In *Statistical Signal Processing Workshop (SSP), 2011 IEEE*, pages 677–680. IEEE, 2011.
- [13] LA Pastur. A simple approach to the global regime of gaussian ensembles of random matrices. *Ukrainian Mathematical Journal*, 57(6):936–966, 2005.
- [14] R. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34(3):276–280, 1986.
- [15] P. Stoica and N. Arye. Music, maximum likelihood, and cramer-rao bound. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 37(5):720–741, 1989.
- [16] P. Vallet, P. Loubaton, and X. Mestre. Improved Subspace Estimation for Multivariate Observations of High Dimension: The Deterministic Signal Case. *IEEE Transactions on Information Theory*, 58(2), Feb. 2012. arXiv: 1002.3234.
- [17] B. Volcker and B. Ottersten. Chirp parameter estimation from a sample covariance matrix. *Signal Processing, IEEE Transactions on*, 49(3):603–612, 2001.